CBP-IDRs regulate acetylation and gene expression

Authors

Katie L. Gelder[1,4,5], Nicola A. Carruthers[1,4,10], Grace Gilbert[1,4,6*], Laura J. Harrison[1,4,8*], Brychan V. Evans[1,4,5], Thomas I. Evans[1], Sophie S. Ball[1,9], Mark Dunning[7], Timothy D. Craggs[3,4,5], Alison E. Twelvetrees[2,4,6], Daniel A. Bose[1,4,5].

Affiliations

1. Molecular and Cellular Biology, School of Biosciences, Faculty of Science, The University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, UK.
2. Division of Neuroscience, The School of Medicine and Population Health, Faculty of Health, The University of Sheffield, Glossop Road, Sheffield, S10 2HQ, UK.
3. Department of Chemistry, School of Mathematical and Physical Sciences, Faculty of Science, The University of Sheffield, Dainton Building, Brook Hill, S3 7HF, UK.
4. Centre for Single Molecule Biology (SM@SH), The University of Sheffield.
5. Nucleic Acids Institute, The University of Sheffield.
6. Neuroscience Institute, The University of Sheffield.
7. Sheffield Bioinformatics Core, Faculty of Health, The University of Sheffield.
8. Current address: Max-Planck-Zentrum für Physik und Medizin, Erlangen.
9. Current address: Department of Immunology and Inflammation, Imperial College London.
10. Current address: MRC Human Genetics Unit, University of Edinburgh, Edinburgh.

*Authors contributed equally to this work.

Corresponding author: Daniel A. Bose, e-mail: d.bose@sheffield.ac.uk

Keywords

CBP/p300, gene regulation, biomolecular condensates, transcription, IDRs, Intrinsically disordered proteins, enhancers, acetylation.

Abstract

Intrinsically disordered regions (IDRs) have emerged as crucial regulators of protein function, allowing proteins to sense and respond to their environment. *Creb binding protein* (CBP) and *EP300* (p300) are transcription coactivators that regulate gene expression in multicellular organisms, following their recruitment to *cis*-regulatory elements. CBP and p300 contain large IDRs, however little is known about how these different IDRs work together to regulate CBP function. Here, we show that CBP-IDRs cooperate to control different aspects of CBP behaviour in the nucleus, by regulating the properties of fluid-like condensates formed by endogenous CBP. We show how IDRs with different sequence properties make unique contributions to CBP behaviour by establishing a balance between positive and negative regulation of CBP condensates. These conflicting interactions are functionally important, shaping CBPs response to factors such as lysine acetylation, that influence condensate formation. When disrupted, regulatory CBP-IDRs change how CBP interacts with chromatin, alter patterns of CBP-dependent histone acetylation and change gene expression. Together, our work highlights how IDRs with different sequences, spatially segregated in the same protein, can cooperate to shape protein function.

Introduction

Intrinsically Disordered Regions (IDRs) are amino acid sequences that regulate many biological functions despite lacking stable secondary and tertiary structures(van der Lee et al. 2014; Holehouse and Kragelund 2023). IDRs mediate specific and non-specific interactions between biomolecules and control the formation and behaviour of biomolecular condensates. Functionally important regions in IDRs can be defined by patterns of amino acid properties or compositional biases[1,2]. Their ability to adopt multiple conformations means that IDRs are responsive to their interactions and environmental context[3,4]. Despite an increasing understanding of the rules which govern these processes, how the different properties of IDRs are combined in the same protein to generate specific functions is still poorly understood.

Creb binding protein (CBP) and its paralog *EP300* (p300) are histone acetyltransferases (HATs) essential for shaping gene expression from *cis*-regulatory elements (CREs). By combining a structured catalytic HAT domain with large IDRs, CBP and p300 integrate HAT activity with the conformational plasticity needed to interact with diverse transcription factors (TFs) at CREs. CBP and p300 finely balance gene transcription by acting in several key ways. They control the strength of transcription by mediating lysine acetylation on histones (e.g., H3K27ac, H3K18ac, H3K122ac)[5–8] and transcription factors (TFs)[9]. Additionally, they contribute to RNA Polymerase II (PolII) recruitment to maintain promoter-proximal pausing[10–12] and form biomolecular condensates to regulate transcriptional bursting[11,13,14]. Intrinsically Disordered Regions (IDRs) and condensates are central to many of these processes. Firstly, binding to TFs is mediated by a combination of structured binding domains and IDRs, many of which undergo 'coupled folding and binding' with TFs [15–20]. Secondly, HAT activity is regulated by the autoregulatory loop (CBP_{AL}), an IDR insertion in the HAT domain that regulates enzymatic activity by blocking substrate binding to the active site[21]. CBP_{AL} undergoes autoacetylation and binds to non-coding enhancer RNAs (eRNAs), both of which displace it from the active site to stimulate activity [21–25] [21,22,26]. p300 HAT activity is decreased by crowding reagents and when condensate formation is triggered *in vitro* [27,28]. Conversely, formation of CBP and p300 condensates driven by the disordered transactivation domains of TFs in cells increased acetylation across the genome[11], while TF binding to regions outside the HAT domain in CBP have been shown to stimulate or inhibit its HAT activity[29]. In combination, this suggests that CBP functions are highly dependent on interactions with its environment, and underlines the importance of IDRs for this process.

Understanding how multiple IDRs with different sequence properties regulate functions when combined in the same protein remains a significant challenge. This is particularly relevant for CBP and p300, where multiple IDRs mediate diverse biological processes [11,19,30]. To address this question, we examined how distinct IDRs within CBP (CBP-IDRs) work together to regulate CBP behaviour, by systematically characterising the impact of these IDRs on CBP's function within cells. We found that different IDRs play very different roles. Two adjacent C-terminal IDRs with different amino acid compositions and sequence patterning together allow CBP to respond to other factors that regulate condensates, such as acetylation. When disrupted, these IDRs caused changes in the behaviour of CBP condensates, including changing how they respond to lysine acetylation, and

altered CBP's chromatin localisation, histone acetylation and gene expression. Our results uncover new and unexpected functional roles for CBP-IDRs with different sequence properties, and highlight how intramolecular cooperation between IDRs can shape protein behaviour.

Results

CBP forms fluid-like condensates in the nucleus.

To validate the formation of CBP condensates, we imaged endogenous CBP by immunofluorescence (IF) in HEK293T cells, revealing distinct nuclear puncta consistent with previous observations[31–34] (Figure 1A). To uncover the behaviour of these puncta, we introduced an in-frame HaloTag to the C-terminus of endogenous CBP (CBP-Halo). PCR from genomic DNA confirmed heterozygous insertion (Figure S1A-B) and we confirmed near-endogenous expression levels by RT-qPCR (Figure S1C) and western blotting (Figure 1B). To confirm that CBP-Halo can also form endogenous puncta, we carried out live-cell imaging using the fluorescent ligand TMR to label the HaloTag (CBP-Halo$_{TMR}$). CBP-Halo$_{TMR}$ puncta were only observable in CBP-Halo HEK293T cells upon addition of TMR (Figure 1C).

Next, we used lattice light sheet microscopy to visualise the spatial organisation of CBP puncta (Figure 1D). CBP-Halo formed an average of 27.9 condensates/nucleus and we observed a variation in the size and integrated intensity density (IID) of observable puncta, with a mean volume of 33 ± 14.6 μm^3 and an IID of 835.6 ± 132.8. Importantly, CBP-Halo$_{TMR}$ condensates demonstrated hallmarks of fluid-like behaviour, as they were able to fuse and disperse over the 2s imaging period in live-cell confocal microscopy (Figure 1E & Supplementary Movie 1). We next carried out Fluorescence Recovery After Photobleaching (FRAP) on endogenous CBP using Janelia Fluor® 549 Halo ligand (CBP-Halo$_{JF549}$), which showed that fluorescent signal recovered to ~50% its pre-bleach levels after 30s, and to ~90% over a 60s timeframe following photobleaching (Figure 1F-G & Supplementary Movie 2). Together, these data clearly establish that endogenous CBP condensates have fluid-like behaviour in their native nuclear environment.

IDRs in CBP have distinct abilities to form condensates.

Full length CBP$_{wt}$ is predicted to be ~65% disordered (Figure 1H)[19]. AlphaFold2 predictions of CBP$_{wt}$ show structured domains interspersed with large disordered regions (Figure S1D)[35]. To test how these regions regulate CBP's function, we identified nine IDRs for further analysis, using a threshold of 50 consecutive amino acids with a predicted disorder >50% (CBP$_{IDR1-IDR7}$, Figure 1H)[36] in PONDR[37] to define IDR boundaries. We also tested the CBP:FUS interaction domain (CBP$_{CFID}$), a region previously shown to bind to Fused-in-sarcoma (FUS)[11,38–41] which comprises two identified IDRs (CBP$_{IDR6}$ and CBP$_{IDR7}$) and the Nuclear coactivator binding domain (CBP$_{NCBD}$), which binds to TFs[15,17]. Finally, we included CBP$_{AL}$, an IDR in the

HAT domain important for regulating CBP activity via acetylation and RNA binding (Figure 1H)[21,22,28,42]. The behaviour of IDRs is governed by non-random amino acid compositions and non-random patterning of amino acids. To uncover compositional and sequence biases in CBP-IDRs, we used NARDINI+, which compares the enrichment and depletion of sequence features of known importance for IDR behaviour, to the distribution of that feature in all human IDR-containing proteins[43,44]. Analysis with NARDINI+ revealed a number of sequence features that were enriched or depleted in specific CBP-IDRs (Figure S1E). For example, CBP_{IDR7} was enriched for glutamine residues and glutamine patches, reflecting the 18-residue polyQ region (residues 2199-2216) and depleted for hydrophobic residue pairs, while CBP_{AL} was enriched for asparagine, lysine patches and positive/negative residue pairs. In contrast, CBP_{IDR4} and CBP_{IDR6} had an elevated ratio of Arginine and Lysine residues, as well as a higher fraction of proline residues and an increased propensity to adopt polyproline II conformations[45], consistent with previous NMR studies[46].

As these factors affect biomolecular condensates [47], we tested the ability of each CBP-IDR to form condensates in HEK293T cells using the OptoDroplet system, which drives condensate formation through homotypic interactions mediated primarily by self-associating IDRs[36,39]. Candidate IDRs and controls were fused to the photoactivatable Photolyase homology region of Cry2 (Cry2PHR) and an SV40 nuclear localization sequence (NLS) to control phase transitions in the nucleus (Figure S1F). Following exposure to blue light, CBP-IDRs with strong condensate forming properties (e.g CBP_{IDR7}) formed defined puncta, while IDRs unable to form condensates (e.g CBP_{IDR6}) failed to form puncta (Figure 1I & S1G). To quantify differences between CBP-IDRs, we measured the standard deviation (SD) of intensities within treated nuclei: as condensate formation causes strong intensity changes, we expected the variance in grey values for a defined nucleus to increase, resulting in a high SD. Notably, most tested CBP-IDRs behaved similarly to the negative control, with the exception of CBP_{IDR7} and CBP_{CFID} which both demonstrated a rapid increase in SD following photoactivation (Figure 1J & S1H). We then calculated the initial rate of droplet formation over the first 30 seconds following photoactivation; again, only CBP_{IDR7} and CBP_{CFID} showed a greater ability to form condensates than the negative control Opto-NLS (Figure 1K & S1I). The results suggest that different CBP-IDRs undergo distinct phase transitions in the nucleus, and highlight strong condensate forming properties in the disordered C-terminal domain of CBP (CBP_{IDR7} and CBP_{CFID}).

CBP-IDRs regulate the number of CBP condensates

We next tested how CBP-IDRs contribute to formation of full-length CBP (CBP_{wt}) condensates. We deleted CBP-IDRs in CBP_{wt} fused to an in-frame C-terminal GFP-tag (CBP_{wt}-GFP, Figure 2A). To generate $CBP_{\Delta AL}$-GFP, we replaced the IDR with a flexible 12-residue linker to maintain correct folding of the HAT domain (Figure 2A)[23]. Western blots show that HEK293T cells expressing CBP-GFP constructs have similar CBP-GFP expression at a population level (Figure S2A), although we cannot rule out the possibility of variable cell-to-cell transfection within each population. Cells were then fixed and imaged by confocal microscopy for quantification of nuclear puncta (Figure 2B).

Data were initially selected for successful transfection based on the presence or absence of GFP signal, then the data was blinded and scored for punctate or diffuse nuclear signal (Figure 2C & S2B). Following expression of CBP_{wt}-GFP, $CBP_{\Delta IDR1}$-GFP or $CBP_{\Delta IDR6}$-GFP, 100% of scored nuclei demonstrated a punctate GFP signal. Indeed, expression of the majority of CBP-GFP constructs resulted in a punctate GFP signal, as $CBP_{\Delta IDR2\text{-}5}$-GFP, and $CBP_{\Delta AL}$-GFP were all >91% punctate. $CBP_{\Delta IDR7}$-GFP and CBP_{CFID}-GFP were exceptions, with only 52.9% and 40.0% of nuclei showing a punctate signal. To provide a more unbiased measure of the proportion of nuclei that formed condensates, we measured the SD of intensities within treated nuclei. Using this approach, $CBP_{\Delta IDR1\text{-}5}$-GFP failed to show any significant change in the formation of punctate nuclei compared to CBP_{wt}-GFP (Figure 2D & S2G). However, $CBP_{\Delta IDR6}$-GFP nuclei were significantly enriched for condensates, while $CBP_{\Delta AL}$-GFP, $CBP_{\Delta IDR7}$-GFP and $CBP_{\Delta CFID}$-GFP were significantly depleted for condensates (Figure 2D). Although informative, this coarse-grained approach was insufficient to describe the full spectrum of condensate-forming $CBP_{\Delta IDR}$-GFP behaviour.

To better understand how CBP-IDRs affect both the number and the size and intensity of condensates, we first applied an unbiased filter to remove oversaturated nuclei, noise and overlapping puncta (Figure S2C-E) and calculated the mean number of condensates per nucleus for each CBP-GFP construct (Figure 2E & S2H). CBP_{wt}-GFP and the majority of CBP-IDR deletions formed well defined condensates, as >80% of nuclei displayed a punctate signal and nuclei that passed the filter formed >6 condensates per nucleus (Figure 2B,E & S2F,H). $CBP_{\Delta IDR1}$-GFP and $CBP_{\Delta IDR3}$-GFP behaved similarly to CBP_{wt}-GFP, while $CBP_{\Delta IDR4}$-GFP and $CBP_{\Delta AL}$-GFP showed a ~50% decrease in the average number of condensates compared to CBP_{wt}-GFP (Figure 2E). There was a high degree of variability in the number of condensates per nucleus (Figure S2F). For CBP_{wt}-GFP, the SD in the number of condensates was 0.86 of the mean, and most CBP-IDR deletions displayed a similarly high ratio, ranging from 1.37 for $CBP_{\Delta IDR1}$-GFP to 0.75 for $CBP_{\Delta IDR7}$-GFP. In contrast to condensate-forming mutants, 60% of $CBP_{\Delta CFID}$-GFP and 47% of $CBP_{\Delta IDR7}$-GFP nuclei displayed a diffuse GFP signal (Figure 2C), alongside a large decrease in the SD of the intensities within each nuclei (Figure 2D) and a ~75% decrease in quantifiable condensates per nucleus compared to CBP_{wt}-GFP (Figure 2E & S2F). Visual inspection of $CBP_{\Delta CFID}$-GFP and $CBP_{\Delta IDR7}$-GFP nuclei showed that unlike other CBP-GFP constructs, any residual condensates which passed our filter were poorly defined puncta (Figure 2B). We were therefore confident that both $CBP_{\Delta CFID}$-GFP and $CBP_{\Delta IDR7}$-GFP prevented condensate formation. Together with the OptoDroplet results (Figure 1 I-K), this emphasised the importance of C-terminal IDRs (CBP_{CFID} and CBP_{IDR7}) for promoting CBP_{wt} condensates. Intriguingly, in contrast to other CBP-IDRs, $CBP_{\Delta IDR6}$-GFP displayed a significantly greater SD in intensity values within each nucleus and formed ~30% more condensates than CBP_{wt}-GFP (Figure 2D & S2F). The variability in the number of condensates for $CBP_{\Delta IDR6}$-GFP was significantly lower (SD was 0.49 of the mean). This result indicates a much more uniform number of condensates compared to CBP_{wt}-GFP and other CBP-IDR deletions, and suggests that the CBP_{IDR6} may actively promote variability in CBP condensate numbers. The results for $CBP_{\Delta IDR6}$-GFP were surprising. Although CBP_{IDR6} (along with CBP_{IDR7}) forms a large part of CBP_{CFID} - which blocked condensation when deleted - the findings suggest that C-terminal IDRs of CBP contain opposing abilities to both promote and inhibit condensate formation.

CBP-IDRs regulate the size and intensity of CBP condensates

Deletion of CBP-IDRs clearly altered the size and intensity of CBP condensates (Figure 2B). To quantify this behaviour, we calculated the Integrated Intensity Density (IID) of CBP condensates. We excluded $CBP_{\Delta CFID}$-GFP and $CBP_{\Delta IDR7}$-GFP which formed too few puncta. CBP_{wt}-GFP condensates had relatively uniform size and intensity (mean IID = 870 ± 72 s.e.m.), and most $CBP_{\Delta IDR}$-GFP constructs again behaved similarly to CBP_{wt}-GFP (Figure 2F & S2I). The two exceptions were $CBP_{\Delta AL}$-GFP, which demonstrated a 1.5-fold increase and $CBP_{\Delta IDR6}$-GFP which demonstrated a 3-fold increase compared to CBP_{wt}-GFP (Figure 2F). The results suggest that when intact, CBP_{IDR6} and CBP_{AL} may restrict the size and intensity of CBP_{wt} condensates.

To confirm that larger CBP-GFP condensates retained fluid-like properties, we carried out FRAP on individual CBP_{wt}-GFP, $CBP_{\Delta AL}$-GFP and $CBP_{\Delta IDR6}$-GFP puncta (Figure 2G-H). Following bleaching, all condensates recovered rapidly (Recovery rate: CBP_{wt}-GFP = 0.11 s^{-1}, $CBP_{\Delta AL}$-GFP = 0.12 s^{-1}, $CBP_{\Delta IDR6}$-GFP = 0.13 s^{-1}, Figure 2H, S2J), demonstrating fluid-like behaviour rather than solid or gel-like properties, and only subtle differences in nuclear mobility.

We also treated cells transfected with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta AL}$-GFP with 5% 1,6-hexanediol (1,6-HD), an aliphatic alcohol that disrupts weak hydrophobic interactions[48]. Condensates formed by all three constructs dissolved in <20s following application of 5% 1,6-HD (Figure 2I), suggesting that hydrophobic interactions are important for maintaining CBP condensates.

Diverse RNAs are important for shaping condensate properties. As CBP binds to RNA *in vitro* and in cells, we therefore tested how it affects CBP condensates (Figure 2I). Treating cells with 0.5uM RNAse A decreased the proportion of punctate nuclei for CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP but not $CBP_{\Delta IDR7}$-GFP (Figure S2K). Notably, the size and intensity of CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP condensates was decreased following addition of RNase A, but residual $CBP_{\Delta IDR7}$-GFP condensates were unchanged. RNase A also caused an increase in the number of condensates formed for all tested constructs (Figure S2L), potentially reflecting a dispersal of condensates following RNA degradation. The results highlight the ability of undefined RNAs to maintain properties of CBP condensates.

Taken together, these results highlight clear differences between condensate behaviour regulated by CBP-IDRs. Notably, CBP condensates are strongly promoted by CBP_{IDR7} and inhibited by CBP_{IDR6}, but when deleted in combination as part of $CBP_{\Delta CFID}$, the phenotype of $CBP_{\Delta IDR7}$ appears dominant. CBP_{AL} and CBP_{IDR6} play a more complex regulatory role, restricting the size and intensity of condensates, with CBP_{IDR6} in particular affecting variability in the number of condensates formed. Large CBP condensates have fluid-like properties and are sensitive to disruption of hydrophobic contacts, and RNAs can help to promote CBP condensate formation.

CBP IDRs regulate endogenous CBP condensates

Endogenous CBP-Halo (Figure 1E-G) and CBP-GFP (Figure 2F-G) both form fluid-like condensates. In cells, condensates form due to a combination of homotypic interactions between self-associating IDRs, and heterotypic multicomponent interactions, including with RNAs and other proteins[49]. We hypothesised that the observed changes in CBP condensate behaviour might be caused, at least partially, by alterations in homotypic interactions between its IDRs. To test whether this was the case, we next asked whether CBP-IDR deletions could regulate endogenous CBP condensates. We used lattice light sheet microscopy to image endogenous CBP-Halo labelled with fluorescent JF_{646} ligand ($CBP\text{-}Halo_{JF646}$) following expression of CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP (Figure 2L). Compared to control cells, CBP_{wt}-GFP caused a slight decrease in the number of endogenous CBP-Halo condensates (27.9 ± 79.2 > 14.9 ± 28.9 s.e.m condensates/nucleus; Figure 2M), coupled to an increase in the IID (835 ± 132 > 1759 ± 803 s.e.m; Figure 2N) and average volume (32.9 ± 14.6 > 58.9 ± 40.0 s.e.m; Figure 2O), probably reflecting the presence of extra CBP-GFP molecules due to transient expression. However, compared to cells expressing CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP increased the median IID (2223 ± 832 s.e.m; Figure 2N) and the volume (101.7 ± 55.4 s.e.m; Figure 2O) of endogenous CBP-Halo condensates, while $CBP_{\Delta IDR7}$-GFP, had the opposite effect, decreasing both the IID (655 ± 411 s.e.m; Figure 2N) and volume (45.2 ± 33.4 s.e.m; Figure 2O). Similar results were seen with live cell confocal microscopy (Figure S2M) with no bleed through into other imaging channels (Figure S2N). Although we cannot rule out complexities arising from the mixed populations of endogenous wild type and co-expressed CBP-GFP, these results highlight the importance of homotypic interactions between CBP-IDRs, with CBP_{IDR6} and CBP_{IDR7} exerting a dominant positive/negatve influence respectively on endogenous CBP condensates. Although striking, the results don't exclude the potential for heterotypic interactions, and indeed we believe this is highly likely, given evidence from p300[11]. The ability of CBP-IDRs to shape condensates through homotypic interactions could provide a pathway for modifications of CBP-IDRs, such as acetylation, or disease mutations in CBP-IDRs, to broadly affect CBP activity by affecting its condensate behaviour.

Opposing properties of CBP-IDRs reflect their sequence characteristics

To investigate whether the opposing behaviours of the C-terminal IDRs CBP_{IDR6} and CBP_{IDR7} are encoded by their amino acid sequence, we used localCIDER[50] to examine their sequence properties, alongside CBP_{CFID} which contains both CBP_{IDR6} and CBP_{IDR7} (Figure 3A). Compared to CBP_{wt}, CBP_{CFID} is enriched for polar amino acids and proline, which promote expanded chain conformations[51], but depleted for charged and aromatic residues. There is more variation between CBP_{IDR6} and CBP_{IDR7} individually: Polar amino acids are more enriched in CBP_{IDR7} (54%) than CBP_{IDR6} (41%), while CBP_{IDR6} is enriched for proline (21%) compared to CBP_{IDR7} (14%). The linear distribution of amino acids (Figure 3B) shows that amino acid properties in CBP_{IDR6} are evenly distributed, but less so in CBP_{IDR7}. This is especially true for polar amino acids, which in CBP_{IDR7}

are concentrated around the 18-residue polyQ region (residues 2199-2216), and proline which is concentrated towards the C-terminus (Figure 3B).

The composition and sequence patterning of amino acids can impart different behaviours to IDR-containing proteins, and create different properties in biomolecular condensates. To test the role of sequence patterning in CBP_{IDR6} and CBP_{IDR7}, we used GOOSE[52] to design CBP_{IDR6} and CBP_{IDR7} mutants where the overall composition of was retained, but their sequence was randomly shuffled ($CBP_{IDR6\text{-}Shuffle}$ and $CBP_{IDR7\text{-}Shuffle}$ Figure S3A-B). Analysis of wild-type and shuffled CBP-IDR sequences using NARDINI+[43,44] highlighted expected similarities in composition, for example in the Fraction of glutamine residues, (FracQ) and Fraction of Charged Residues (FCR) (Figure 3C, bold). It also highlighted differences in pairwise patterning in the shuffled sequences; for example, CBP_{IDR6} was enriched for leucine patches and pairs of hydrophobic-proline and proline-proline residue pairs, while $CBP_{IDR6\text{-}shuffle}$ was depleted for these features. Similarly, while CBP_{IDR7} was depleted for hydrophobic and polar-proline residue pairs, $CBP_{IDR7\text{-}shuffle}$ was enriched for these features (Figure 3C, bold).

Next, we tested whether these differences in sequence patterning were important for the formation of CBP condensates. We expressed $CBP_{IDR6\text{-}Shuffle}$-GFP and $CBP_{IDR7\text{-}Shuffle}$-GFP in HEK293T cells and imaged using confocal microscopy (Figure 3D). $CBP_{IDR6\text{-}Shuffle}$-GFP formed large condensates, similar to those formed by $CBP_{\Delta IDR6}$-GFP. Following filtering (Figure S3D), it was clear that shuffling the residues in CBP_{IDR6} increased the proportion of punctate nuclei (Figure S3E), the number of condensates per nuclei (Figure 3E) and the size and intensity of condensates compared to CBP_{wt}-GFP (Figure 3F), while condensates became more uniform. The result suggests that the order and sequence patterning of amino acids in CBP_{IDR6} is important for its condensate forming behaviour, as disrupting the native sequence has the same effect on phenotype as deleting the IDR. In contrast, $CBP_{IDR7\text{-}Shuffle}$-GFP continued to form condensates, unlike $CBP_{\Delta IDR7}$-GFP. Indeed $CBP_{IDR7\text{-}Shuffle}$-GFP behaved more similarly to CBP_{wt}-GFP, with no difference in the proportion of punctate nuclei (Figure S3E), number of condensates per nuclei (Figure 3D) or the size and intensity of condensates (Figure 3E). The result suggests that the ability of CBP_{IDR7} to drive condensate formation is more likely to result from its overall composition - which includes enrichment of glutamine and the ratio of arginine and lysine residues (Figure 3C, bold) - than specific pairwise patterning of amino acids.

Our results show that CBP_{IDR7} strongly promotes CBP condensates while CBP_{IDR6} restricts condensate formation. CBP_{IDR6} and CBP_{IDR7} are enriched for polar residues and polyQ tracts (Figure 3C), which are associated with prion-forming domains[53]. Prion-like domains have been identified as strong drivers of condensate formation across diverse proteins[54]. To test whether CBP_{IDR6} and CBP_{IDR7} could be displaying prion-like behaviours, we used PLAAC analysis[53,55] to examine the sequence of the C-terminal CBP-IDRs for prion-like regions. Results from PLAAC predicted that CBP_{IDR7} forms a large, contiguous prion-like domain (Figure 3F), which extended even further towards the C-terminus of CBP in $CBP_{IDR7\text{-}Shuffle}$. This is consistent with the ability of CBP_{IDR7} to form condensates in OptoDroplet experiments (Figure 1I-K), and the inability of $CBP_{\Delta IDR7}$-GFP to form condensates (Figure 2B-E). While CBP_{IDR6} was also predicted to contain prion-like regions, these were discontiguous and smaller than those encoded by CBP_{IDR7}.

Finally, we also tested the role of residues 1951-1978 in CBP_{IDR6} ($CBP_{IDR6-H2}$), which were shown to form transient α-helices by Nuclear Magnetic Resonance (NMR)[46], and were predicted by Alphafold2 to interact with an α-helix formed by residues 2187-2216 in CBP_{IDR7} (Figure S3F, $CBP_{IDR7-H1}$). Deletion of either predicted helix did not affect CBP condensates (Figure S3G-I).

The different behaviours caused by CBP_{IDR6} and CBP_{IDR7} led us to ask whether their sequence properties were conserved in CBP's paralog p300. Over their entire sequence, CBP and p300 are 60.8% identical (70.7% similarity). While this identity is as high as 87.3% (94.4% similarity) in the HAT domain, CBP_{IDR6} only has 48.1% identity (58.0% similarity) and CBP_{IDR7} has 56.1% identity (66.2% similarity) (Figure S3J). To compare the sequence properties of CBP_{IDR6} and CBP_{IDR7} to their counterparts in p300, we again used NARDINI+[43,44] (Figure S3K). The results highlight different composition and patterning parameters between $p300_{IDR6}$ and $p300_{IDR7}$ as seen for CBP, however there were also notable divergences between the equivalent sequences in CBP and p300. For example, CBP_{IDR7} has higher enrichment of polar-polar and polar-glycine residue pairs, while $p300_{IDR7}$ has a higher ratio of arginine and lysine residues and is more depleted for polar-hydrophobic residue pairs (Figure S3K, bold). As CBP_{IDR6} and CBP_{IDR7} were predicted to contain prion-forming domains, we also subjected p300 to PLAAC analysis. Similar to CBP, $p300_{IDR7}$ was predicted to form a large, contiguous prion-like domain. However, where prion-like properties in CBP_{IDR6} were discontinuous, $p300_{IDR6}$ was predicted to also form a single prion-like domain. Taken together, results based solely on their amino acid sequences suggest that condensates formed by p300 may behave differently to those formed by CBP, but further work is needed to test the hypothesis experimentally.

CBP-IDRs allow global lysine acetylation to shape CBP condensates.

Our results show that distinct sequence properties within CBP-IDRs lead to markedly different properties in the resulting CBP condensates. We hypothesised that this could influence how other environmental inputs, such as post-translational modifications (PTMs), control condensate behaviour and affect overall CBP function. To test this, we focused on lysine acetylation: CBP and p300 acetylate nearly one-third of nuclear proteins, including themselves[9], and thus both are acetylated on multiple residues, including a hotspot located in CBP_{AL}[9,21,28]. Acetylation therefore plays a complex role in regulating CBP/p300 function. Recently, acetylation has been shown to regulate p300 condensates in cells by providing nucleation sites for bromodomain proteins,[11,28,56] and *in vitro* where recombinant p300-HAT domain condensates are promoted by $p300_{AL}$'s positive charge and negatively regulated by lysine acetylation[28]. However, it remains unknown how these regulatory events combine with CBP-IDR-mediated regulation of condensates.

Firstly, to test how CBP-IDRs allow CBP to respond to global lysine acetylation, we used A-485, a potent inhibitor of CBP and p300 HAT activity[57] to deplete CBP and p300-dependent acetylation and asked how condensates formed by $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP behaved (Figure 4A). Treatment with 500nM A-485 for 36 hours[57] increased the proportion of punctate nuclei in CBP_{wt}-GFP (Figure 4B), but had no effect on the

number (Figure 4C) or intensity of condensates (Figure 4D). In contrast, A-485 caused a slight decrease in both the proportion of $CBP_{\Delta IDR6}$-GFP cells forming condensates and the number of puncta per nuclei (Figure 4B-C), and a large decrease in the size and intensity of condensates (Figure 4D). For $CBP_{\Delta IDR7}$-GFP, A-485 caused more cells to form puncta (Figure 4B), but when quantified the size and intensity of condensates was lower than control cells (Figure 4D). The results confirm that CBP and p300-dependent acetylation can regulate CBP condensates. The large decrease in the size and intensity of $CBP_{\Delta IDR6}$-GFP condensates suggests that removal of CBP_{IDR6} increased their sensitivity to changes in global acetylation, while acetylation was unable to overcome the dramatic decrease in condensate formation caused by deletion of CBP_{IDR7}.

We next carried out the opposite experiment, testing the response of condensates to hyperacetylation and chromatin de-compaction by treating cells with Trichostatin A (TSA), a general inhibitor of histone deacetylases (HDACs). Treatment with 500nM Trichostatin A (TSA)[58] for 90 minutes caused a slight increase in the proportion of punctate nuclei in CBP_{wt}-GFP and $CBP_{\Delta IDR7}$-GFP cells (Figure 4B), but had a negligible effect on the number and integrated density of condensates. However, following TSA treatment, $CBP_{\Delta IDR6}$-GFP showed an increase in both the number of condensates per nuclei and the size and intensity of condensates (Figure 4B-D), suggesting that acetylation could positively regulate $CBP_{\Delta IDR6}$-GFP condensates. Again this suggests that CBP has increased sensitivity to acetylation following removal of CBP_{IDR6}.

We were surprised by the A-485 data for CBP_{wt}-GFP, as A-485 was previously shown to cause a large increase in the number, but decrease the size of $p300_{wt}$-GFP puncta when used at a higher concentration[11]. We therefore tested how CBP condensates responded to different concentrations (including 10µM) of A-485 and shorter treatment times (0-120 minutes)[11] (Figure 4E & S4C). Consistent with our initial data, increasing the concentration of A-485 for 60 and 120 minutes did not change the number of CBP_{wt}-GFP puncta (Figure 4F & S4D, orange), and slightly decreased their intensity, which only reached significance after treatment with 10uM A-485 for 120 minutes (Figure 4G & S4E, orange). In contrast, increasing concentrations of A-485 caused a pronounced decrease in the number (Figure 4E & S4D, blue) of $CBP_{\Delta IDR6}$-GFP condensates, supporting an increased sensitivity to A-485. However, while 60 minutes treatment with 10µM A-485 caused a modest decrease in intensity (Figure 4G, blue), we note that no changes in intensity were observed after 120 minutes treatment (Figure S4E, blue).

Experiments with A-485 and TSA reveal the potential importance of CBP_{IDR6} for shaping how CBP condensates behave in response to global acetylation. Its removal appears to hyper-sensitize CBP condensates to changes in acetylation, as hypoacetylation caused a greater decrease, while hyperacetylation caused a greater increase in the number, size and intensity of condensates than CBP_{wt}-GFP. However, A-485 and TSA both affect global levels of acetylation, which makes it challenging to separate changes in condensates due to direct acetylation of CBP from indirect disruption of heterotypic interactions.

Regulation of condensates by acetylation of CBP_{AL} relies on CBP-IDR integrity.

Next, to assess the role of direct CBP acetylation on condensate formation, we examined how changing acetylation of CBP_{AL} regulates CBP condensates. In $CBP_{\Delta AL}$-GFP data, we observed a ~50% reduction in the mean number of condensates (Figure 2E) but a slight increase in their size and intensity (Figure 2F) compared to CBP_{wt}-GFP. This might be explained by a balance in CBP condensates: a positive effect from exposed lysines in CBP_{AL} could be offset by negative regulation due to lysine acetylation, as seen previously for recombinant p300-HAT *in vitro*[28]. To test the role of acetylation on CBP_{AL}, we generated two mutants, $CBP_{\Delta AL}$-GFP (Figure 2A) and CBP_{KTG}-GFP, where all 13 lysine residues were mutated to glycine to prevent acetylation (Figure S4F). When expressed in HEK293T cells to visualise CBP condensates, $CBP_{wt/\Delta AL}$-GFP and $CBP_{wt/KTG}$-GFP had similar effects, decreasing the overall proportion of punctate nuclei compared to CBP_{wt}-GFP (Figure 4H & 4I). Both CBP_{AL} mutations also reduced the number of condensates per nucleus by 40-60% compared to CBP_{wt}-GFP, but retained a high variability (Figure 4J) and caused a 40-50% increase in their size and intensity (Figure 4K). As CBP_{AL} in $CBP_{wt/KTG}$-GFP was intact, but incapable of being acetylated, this suggested that acetylation of lysine residues in CBP_{AL} can help to maintain a larger number of smaller CBP condensates.

To better understand the relative contributions of CBP-IDRs and lysine acetylation on the CBP_{AL}, we next tested how CBP_{IDR6} and CBP_{IDR7} impacted condensate behaviour due to CBP_{AL} acetylation. To do this, we made double mutants of CBP-GFP, where the two single CBP_{AL} mutants ($CBP_{\Delta AL}$ & CBP_{KTG}) were generated alongside $CBP_{\Delta IDR6}$-GFP ($CBP_{\Delta IDR6/\Delta AL}$-GFP & $CBP_{\Delta IDR6/KTG}$-GFP) and $CBP_{\Delta IDR7}$-GFP ($CBP_{\Delta IDR7/\Delta AL}$-GFP & $CBP_{\Delta IDR7/KTG}$-GFP). Due to their inability to form condensates, $CBP_{\Delta IDR7}$-GFP and $CBP_{\Delta CFID}$-GFP were not affected by CBP_{AL} mutations (Figure S4G-I). Notably, CBP_{AL} mutations did not produce the same effects when combined with $CBP_{\Delta IDR6}$-GFP as they did in CBP_{wt}-GFP. When combined with $CBP_{\Delta IDR6}$-GFP, neither CBP_{AL} mutation affected the overall proportion (Figure 4I), the number (Figure 4J) or the size and intensity (Figure 4L) of condensates compared to $CBP_{\Delta IDR6}$-GFP single mutants. The result suggests that $CBP_{\Delta IDR6}$-GFP is essential for enabling these condensates to respond to acetylation, as mutations in CBP_{AL} that alter CBP_{wt}-GFP condensate behaviour lost their effect when $CBP_{\Delta IDR6}$-GFP was deleted.

Finally, to compare how direct acetylation of CBP_{AL} works together with global acetylation, we treated cells expressing $CBP_{wt/KTG}$-GFP or $CBP_{\Delta IDR6/KTG}$-GFP with A-485. Following treatment, there was a slight reduction in the size and intensity of $CBP_{wt/KTG}$-GFP, and a larger decrease in $CBP_{\Delta IDR6/KTG}$-GFP condensates (Figure S4J) as observed following treatment of CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP with A-485. As neither mutant could be acetylated on CBP_{AL}, the result underlines the importance of bulk acetylation as a driver of condensates, and again highlights how deletion of CBP_{IDR6} appears to hypersensitize condensates to changes in global acetylation levels.

394 Together, these results highlight the complex roles of acetylation on CBP condensates. They suggest that
395 changes to global acetylation exert relatively minor effects on CBP_{wt} condensates, while direct acetylation of
396 CBP_{AL} could support formation of more and smaller condensates. The results clearly demonstrate the
397 importance of CBP IDRs; in all cases deletion of CBP_{IDR7} severely reduced the ability of CBP to form
398 condensates, while deletion of CBP_{IDR6} promoted the formation of more and larger condensates regardless of
399 mutations that affect acetylation on CBP_{AL}. This latter finding suggests that the integrity of CBP_{IDR6} could be
400 crucial to ensure that modifications, such as acetylation on CBP_{AL}, can effectively regulate condensates.
401 However, observations made using CBP_{AL} mutations must be interpreted cautiously before conclusively
402 attributing results to lysine acetylation. Due to the importance of CBP_{AL} for CBP function, these mutations
403 could have multiple pleiotropic consequences, each capable of shaping CBP condensate behaviour. For
404 example, deletion of CBP_{AL} increases catalytic activity[21] which could lead to increased global histone
405 acetylation and nucleation of condensates[56]. Similarly, lysine residues in CBP_{AL} facilitate RNA binding[22],
406 potentially affecting condensates independent of acetylation[59].

407

408 IDRs regulate CBP-dependent acetylation

409 Next, we asked whether CBP-IDRs could affect CBP-dependent histone acetylation. To better control
410 expression of CBP-GFP, we made clonal HEK293T lines stably expressing doxycycline inducible CBP_{wt}-GFP,
411 $CBP_{\Delta IDR6}$-GFP and CBP_{IDR7}-GFP. The cells showed a clear increase in GFP signal in response to increasing
412 concentrations of doxycycline (Figure S4K-L). For all subsequent experiments we chose doxycycline
413 concentrations that maintained overall CBP-GFP expression at near-endogenous levels (Figure S4L).

414 Following induction, CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP formed visible condensates, while $CBP_{\Delta IDR7}$-GFP
415 displayed a more diffuse signal (Figure 4M). The overall proportion of punctate nuclei (Figure S4M) and the
416 number of condensates per nuclei was lower for inducible CBP-GFP than when transiently expressed (Figure
417 S4N). $CBP_{\Delta IDR6}$-GFP condensates (mean IID = 126.0 ± 21.5 s.e.m.) were also larger and more intense than
418 CBP_{wt}-GFP condensates (mean IID = 80.6 ± 18.3 s.e.m.), while $CBP_{\Delta IDR7}$-GFP failed to form quantifiable
419 puncta (Figure S4O).

420 Next, we tested how CBP-IDRs impacted histone acetylation. Induction of CBP_{wt}-GFP slightly increased total
421 levels of H3K27ac, while $CBP_{\Delta IDR7}$-GFP slightly decreased global H3K27ac (Figure 4N & S4P). However,
422 induction of $CBP_{\Delta IDR6}$-GFP caused >2-fold increase in total H3K27ac (Figure 4N & S4P). The amount of
423 activated CBP (K1535ac) remained similar relative to total CBP levels for CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP, but
424 was depleted by $CBP_{\Delta IDR7}$-GFP (Figure 4N & S4Q). To test whether HAT activity was affected by
425 $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP, we purified doxycycline inducible CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and
426 $CBP_{\Delta IDR7}$-GFP and tested their ability to acetylate recombinant nucleosomes *in vitro*. We expected that dilution
427 of CBP and disruption of chromatin during sample preparation would prevent formation of condensates *in*
428 *vitro*. CBP_{wt}-GFP had the highest HAT activity, while surprisingly, H3K27ac was reduced for $CBP_{\Delta IDR6}$-GFP

429 and further decreased for $CBP_{\Delta IDR7}$-GFP (Figure 4O & S4R). The disparity between the total level of H3K27ac
430 in cells, which increased following induction of $CBP_{\Delta IDR6}$-GFP, and the decreased activity of purified
431 $CBP_{\Delta IDR6}$-GFP suggests a complex regulatory mechanism. The total level of CBP (and p300)-dependent
432 acetylation observed in cells is a function of both the acetyltransferase activity of CBP, and the recruitment of
433 CBP to sites on chromatin. The discrepancy could therefore be explained by an increase in the contacts
434 between CBP and chromatin as a result of the formation of larger condensates, or by a stimulation of
435 $CBP_{\Delta IDR6}$-GFP activity in cells that we are unable to detect *in vitro*.

436 CBP IDRs shape how CBP binds to chromatin.

437 Our results suggest that CBP-IDRs outside the HAT domain can influence HAT activity, and highlight a
438 discrepancy between CBP behaviour *in vitro* and levels of histone acetylation in cells. We reasoned that as
439 well as controlling HAT activity, regulation of CBP condensates by CBP-IDRs could change CBP chromatin
440 association, and thus control total levels of histone acetylation. We tested this in our inducible CBP-GFP cell
441 lines using chromatin immunoprecipitation sequencing (ChIP-seq) and CUT&RUN.

442 To confirm the GFP-tag does not adversely affect CBP recruitment to chromatin, we first verified that induced
443 CBP-GFP binds to the same genomic locations as endogenous CBP. We compared ChIP-seq profiles for
444 endogenous CBP_{wt} in maternal $HEK293T_{Tet3G}$ cells (α-CBP; $CBP_{wt\text{-}Tet3G}$) and for CBP-GFP (α-GFP; following
445 addition of Doxycycline to inducible cell lines (Figure 4L). Read densities for CBP ChIP-seq ($CBP_{wt\text{-}Tet3G}$) and
446 GFP ChIP-seq were well correlated across the genome (Figure S5A&B), confirming that reads aligned to
447 similar locations. We then used MACS2[60] to highlight CBP-bound sites in pooled replicates (Figure 5A),
448 identifying 14,913 binding sites for endogenous $CBP_{wt\text{-}Tet3G}$ using CBP ChIP-seq, but only 7,661 sites for
449 CBP_{wt}-GFP in GFP ChIP-seq (Figure 5A). We attribute this disparity to reduced precipitation in the GFP
450 ChIP-seq data, likely stemming from differences in antibody efficiency and from the mixed population of
451 endogenous CBP and CBP-GFP proteins. Accordingly, while endogenous $CBP_{wt\text{-}Tet3G}$ binding sites were still
452 enriched for CBP_{wt}-GFP read densities in GFP ChIP-seq data, the enrichment was lower than that observed
453 with endogenous $CBP_{wt\text{-}Tet3G}$ read densities from CBP ChIP-seq data (Figure S5C-D); endogenous $CBP_{wt\text{-}Tet3G}$
454 sites (median width = 416bp) were therefore broader than those of CBP_{wt}-GFP (median width = 340bp, Figure
455 5B), again indicative of differences in ChIP efficiency. Despite differences in the number of binding sites,
456 CBP_{wt}-GFP and $CBP_{wt\text{-}Tet3G}$ bound to similar regions in the genome; most binding sites were found at promoters
457 (51.5% CBP_{wt}-GFP; 52.1% $CBP_{wt\text{-}Tet3G}$), then introns (14.4% CBP_{wt}-GFP; 18.6% $CBP_{wt\text{-}Tet3G}$) and intergenic
458 sites (13.5% CBP_{wt}-GFP; 15.0% $CBP_{wt\text{-}Tet3G}$); 20.5% of CBP_{wt}-GFP binding sites and 14.3% of $CBP_{wt\text{-}Tet3G}$
459 binding sites mapped to exons (Figure 5C). $CBP_{wt\text{-}Tet3G}$ and CBP-GFP binding sites were closer than expected
460 by chance[61] (Figure S5E-F), demonstrating spatial correlation between CBP-GFP and CBP_{wt}; all CBP-GFP
461 binding sites showed statistically-significant overlap with $CBP_{wt\text{-}Tet3G}$ (Figure 5A & S5F). Taken together, the
462 data show that CBP-GFP and endogenous $CBP_{wt\text{-}Tet3G}$ bind to similar sites in the genome.

463

Given the differences in ChIP efficiency between endogenous CBP and CBP-GFP ChIP-seq datasets, we focused on CBP-GFP ChIP-seq data to uncover how CBP-IDRs affected chromatin localisation. Notably, we observed more binding sites for both $CBP_{\Delta IDR6}$-GFP (18,357 sites) and $CBP_{\Delta IDR7}$-GFP (17,237 sites) than for CBP_{wt}-GFP (7,661 sites) in GFP ChIP-seq (Figure 5A) and the location of binding sites shifted towards introns and intergenic regions for $CBP_{\Delta IDR6}$-GFP (22.6 % introns, 19.9 % intergenic) and $CBP_{\Delta IDR7}$-GFP (27.2 % introns, 21.8% intergenic), (Figure 5C). CBP_{wt}-GFP sites overlapped 77% of $CBP_{\Delta IDR6}$-GFP and 69% of $CBP_{\Delta IDR7}$-GFP. However, only 43% of $CBP_{\Delta IDR7}$-GFP bound-sites overlapped $CBP_{\Delta IDR6}$-GFP (Figure 5A & S5F).

We next asked how CBP-IDRs affected the profiles of CBP-GFP chromatin binding. Compared to CBP_{wt}-GFP at endogenous CBP binding sites, $CBP_{\Delta IDR6}$-GFP read densities were enriched, while $CBP_{\Delta IDR7}$-GFP was depleted (Figure S5C&D). Increased CBP enrichment was reflected in the size of peaks; across all CBP-GFP binding sites, the median peak width for $CBP_{\Delta IDR6}$-GFP (394bp) was greater than for CBP_{wt}-GFP (340bp), and the peak width for $CBP_{\Delta IDR7}$-GFP (283bp) was lower than for CBP_{wt}-GFP (Figure 5B). The behaviour of CBP-GFP constructs when bound to chromatin thus mirrors their behaviour in imaging data: $CBP_{\Delta IDR6}$-GFP bound to more locations with a larger average peak size, while $CBP_{\Delta IDR7}$-GFP bound to more locations but with a smaller average peak size. The results suggest that larger condensates formed by $CBP_{\Delta IDR6}$-GFP increased its enrichment on chromatin, while dispersion of $CBP_{\Delta IDR7}$-GFP condensates led to lower levels of CBP binding, but at a larger number of locations.

CBP IDRs shape histone acetylation.

Our results show that CBP-IDRs regulate total levels of CBP-dependent H3K27ac (Figure 4N) and CBP activity (Figure 4O). To examine how changes in chromatin localisation affected histone acetylation across the genome, we used spike-in normalised CUT&RUN to quantify H3K27ac levels. Genome browser views of CBP-bound regions and corresponding H3K27ac enrichment, such as at the *TSC22D3* gene and a candidate CRE (cCRE) (Figure 5D), highlight binding of CBP-GFP to similar locations as CBP_{wt} and corresponding enrichment for H3K27ac. We used MACS3[60] to identify enriched H3K27ac domains in our CUT&RUN data, highlighting 37,117 loci with a median peak width of 381bp following CBP_{wt}-GFP expression (Figure 5E). Deleting CBP-IDRs had a striking effect on H3K27ac enrichment; compared to CBP_{wt}-GFP, expression of $CBP_{\Delta IDR6}$-GFP caused an increase in the number (46,576 enriched loci) and the size (median width = 421bp) of enriched H3K27ac domains, while expression of $CBP_{\Delta IDR7}$-GFP caused a large decrease in their number (9,213 loci) but not their size (median width = 420bp) (Figure 5E). This corresponded to an increase in both H3K27ac (Figure 5F) and CBP-GFP enrichment compared to CBP_{wt}-GFP (Figure 5G) in $CBP_{\Delta IDR6}$-GFP cells, and a decrease in $CBP_{\Delta IDR7}$-GFP cells. The pattern of H3K27ac and CBP-GFP enrichment was replicated when sampling all CBP-GFP bound sites in CBP-GFP ChIP-seq data (Figure S5G-H).

498 We noted that $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP were recruited to more than twice as many loci as CBP_{wt}-GFP
499 (Figure 5A). To directly compare how CBP-IDRs affect H3K27ac levels, and avoid complexities from
500 sampling a mixture of CBP-bound and CBP-unbound locations, we focussed on a subset of 4,765 binding sites
501 found in common between all CBP-GFP datasets (Figure 5A, magenta). Again, these common sites were
502 enriched for H3K27ac read densities in $CBP_{\Delta IDR6}$-GFP cells, but depleted in $CBP_{\Delta IDR7}$-GFP cells compared to
503 CBP_{wt}-GFP (Figure 5H & S5I). Compared to CBP_{wt}-GFP (mean signal = 0.17), CBP-GFP read densities were
504 enriched at common sites in both $CBP_{\Delta IDR6}$-GFP cells (mean signal = 0.21) and $CBP_{\Delta IDR7}$-GFP cells (mean
505 signal = 0.18) (Figure 5I & S5I). Although CBP-GFP data was not normalised to spike-in controls, the result
506 suggests that H3K27ac broadly reflects CBP recruitment for CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP, but suggests a
507 discrepancy between $CBP_{\Delta IDR7}$-GFP recruitment and H3K27ac.

508 We next mapped the genomic distribution of the unique binding sites identified for each CBP-GFP mutant.
509 Common binding sites predominantly overlapped promoters (58%) and exons (22%), while introns (11%) and
510 intergenic regions (9%) were less common (Figure 5J). This balance shifted considerably for unique sites, with
511 unique sites for CBP_{wt}-GFP (21% introns, 20% intergenic), $CBP_{\Delta IDR6}$-GFP (27% introns, 23% intergenic) and
512 $CBP_{\Delta IDR7}$-GFP (33% intron, 27% intergenic) for frequently found in introns and intergenic regions. This
513 change in distribution suggests that both the larger condensates formed by $CBP_{\Delta IDR6}$-GFP, and the dispersal of
514 $CBP_{\Delta IDR7}$-GFP condensates caused a relocation of CBP to introns and intergenic sites. $CBP_{\Delta IDR6}$-GFP and
515 $CBP_{\Delta IDR7}$-GFP read densities were enriched at their unique binding sites compared to CBP_{wt}-GFP (Figure 5K
516 & S5J) and levels of H3K27ac were elevated at unique $CBP_{\Delta IDR6}$-GFP sites and depleted at unique
517 $CBP_{\Delta IDR7}$-GFP sites (Figure 5L & S5K).

518 Finally, as CBP was recruited to CREs across the genome, we examined how CBP-IDRs affected CBP
519 recruitment and H3K27ac levels at candidate CREs (cCREs) defined by ENCODE[62]. Compared to CBP_{wt}-GFP,
520 binding of $CBP_{\Delta IDR6}$-GFP was enriched and $CBP_{\Delta IDR7}$-GFP were reduced at cCREs (Figure 5M). H3K27ac had
521 a similar pattern, showing enrichment in $CBP_{\Delta IDR6}$-GFP cells and depletion in $CBP_{\Delta IDR7}$-GFP cells, compared
522 to CBP_{wt}-GFP (Figure 5N).

523 Taken together, the data highlight how the effect of CBP-IDRs on HAT activity and chromatin localisation
524 combine to determine patterns of histone acetylation at CBP binding sites on chromatin, and that this
525 behaviour is also seen at cCREs across the genome.

526 CBP-IDRs affect gene expression

527 Our results show that CBP-IDRs regulate the properties of CBP condensates, and thus shape the localisation of
528 CBP to chromatin and the pattern of histone acetylation at regulatory elements across the genome. As CBP
529 binding to CREs and altered histone acetylation directly regulate gene expression, we therefore asked whether
530 disrupting CBP-IDRs affected patterns of gene expression. We carried out RNAseq from two biological
531 replicates each of CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP following addition of doxycycline to induce
532 their expression (Figure 4M). Replicates from the same sample type were well correlated (Figure S6A), and

533 replicates from each sample formed distinct clusters in principle components analysis (PCA) (Figure S6B). To
534 identify gene expression changes specifically resulting from disruption of CBP-IDRs rather than induced
535 expression of CBP-GFP, we compared gene expression profiles in $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP to
536 CBP_{wt}-GFP. Expression of $CBP_{\Delta IDR6}$-GFP resulted in significant up regulation of 190 genes, and down
537 regulation of 349 genes compared to CBP_{wt}-GFP (Figure 6A & S6C); expression of $CBP_{\Delta IDR7}$-GFP caused
538 upregulation of 182 genes and down regulation of 278 genes compared to CBP_{wt}-GFP (Figure 6B & S6D).

539 Importantly, genes with altered expression displayed differences in CBP chromatin binding in GFP ChIP-seq
540 data, and were characterised by changes to H3K27ac levels in spike-in normalised CUT&RUN data. For
541 example *HAND1*, which was upregulated in both $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP datasets, showed
542 additional CBP-GFP binding at regulatory elements, including the promoter, and altered patterns of H3K27ac
543 (Figure 6C). The results highlight how disruption of CBP-IDRs causes off-target accumulation of CBP and
544 alterations to H3K27ac, leading to widespread deregulation of gene expression.

545

546 Discussion

547 Our results provide direct evidence for the importance of CBP-IDRs as regulators of CBP function. We show
548 that different CBP-IDRs can regulate CBP activity in a sequence-dependent manner, by controlling the
549 properties of CBP condensates, altering chromatin localisation and by shaping histone acetylation and gene
550 expression profiles. Importantly, we uncover very different behaviours in different CBP-IDRs, and show how
551 they regulate the response of CBP to lysine acetylation.

552 The ability of condensates to regulate transcription requires balance: too much condensation downregulates
553 transcription, suggesting that condensates must exist within a window of optimal properties to guarantee
554 accurate transcriptional output[11,13,47,48,63]. Here, we highlight how different IDRs within CBP help maintain this
555 balance. Primary drivers of CBP condensate formation are found in its C-terminal IDRs, where CBP_{IDR6} and
556 CBP_{IDR7} make opposing positive and negative contributions to the formation of CBP condensates, that depend
557 on their amino acid sequence patterning in the case of CBP_{IDR6} and overall composition in the case of CBP_{IDR7}.
558 The inhibitory effect of CBP_{IDR6} on the condensates driven by CBP_{IDR7} (Figure 7A) is similar to regulatory
559 interactions in those formed by Ataxin-2, where condensate forming IDRs - with prion-like properties similar
560 to CBP_{IDR7} - could be quenched by neighbouring IDRs[64]. In CBP, the balance between CBP_{IDR6} and CBP_{IDR7}
561 strongly influences how CBP interacts with chromatin. CBP and p300 are recruited to chromatin by the
562 association of TFs with TF interacting domains in their C-terminal regions[65]. Here, we show that disruption of
563 CBP_{IDR6} and CBP_{IDR7} can shape how CBP behaves once it has been recruited to chromatin, potentially by
564 changing its interactions with TF-condensates[11]. This leads to two key changes: an increase in the number of
565 CBP binding sites on chromatin, and altered profiles of CBP recruitment. Specifically, broader interactions
566 occur when CBP_{IDR6} is deleted, while the size of CBP binding sites decreases following CBP_{IDR7} deletion,
567 behaviour which directly mirrors that observed in CBP condensates when CBP_{IDR6} and CBP_{IDR7} are deleted.

These alterations in chromatin binding also reshape histone acetylation patterns at regulatory elements, impacting gene expression. Since CBP functions as a crucial scaffolding factor, modifications to its condensate properties and chromatin binding could further disrupt the recruitment and retention of other vital regulatory factors at these sites[10,12,14].

We have shown that CBP-IDRs determine the sensitivity of CBP condensates to different forms of lysine acetylation. p300 condensates are regulated by global lysine acetylation, which nucleates condensate assembly by recruiting bromodomain proteins[11,56], but also through direct acetylation of lysine resides in p300, which is inhibitory to p300 HAT domain condensates *in vitro*[28]. In contrast to p300, CBP condensates appear relatively insensitive to changes in global acetylation (Figure 7B). This buffering function is potentially mediated by CBP_{IDR6}, as its deletion seems to hyper-sensitize CBP condensates to changes in global acetylation. Like p300, acetylation of CBP_{AL} appears to negatively regulate the number and size of CBP condensates, but again this function is dependent on the integrity of C-terminal CBP-IDRs; deletion of CBP_{IDR7} largely inhibits CBP condensates, while the larger condensates formed following removal of CBP_{IDR6} seem to be largely impervious to acetylation events on CBP_{AL} (Figure 7B). Taken together, the results highlight the fundamental importance of CBP-IDRs to condensate behaviour, but also underline the complex nature of CBP condensates, which integrate diverse environmental factors including global and direct lysine acetylation on CBP.

Surprisingly, as well as CBP's response to acetylation, disrupting CBP-IDRs changed CBP's catalytic activity. CBP-HAT activity was eliminated by deletion of CBP_{IDR7} and drastically reduced by deletion of CBP_{IDR6} (Figure 4O). Truncation of p300 in the $p300_{TAZ2}$ domain (N-terminal to CBP_{IDR6} and CBP_{IDR7}) can increase histone acetylation *in vitro* by relieving $p300_{TAZ2}$-dependent inhibition of HAT activity[66]. As CBP_{TAZ2} remained intact in both our $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP mutants, it's possible that CBP_{IDR6} and CBP_{IDR7} may relieve inhibitory interactions between CBP_{TAZ2} and CBP-HAT, and that their removal causes inhibition of HAT activity by CBP_{TAZ2}. The regulation of HAT activity by C-terminal CBP-IDRs suggests that factors interacting with these regions, such as the CBP_{NCBD} - which is located between CBP_{IDR6} and CBP_{IDR7} - may regulate activity by influencing CBP-IDR behaviour. For example, the adenovirus protein Early region 1A (E1A) can bind to multiple sites in the N- and C-terminal tails of CBP, but is only able to inhibit the HAT activity of CBP when it interacts with a region overlapping CBP_{IDR6}. Mutations in E1A which specifically disrupt its binding to CBP_{IDR6}, but not other sites in CBP[29], resulted in the loss of this inhibition. Similarly, stimulation of CBP HAT activity was seen with mutants of Activator of Thyroid and Retinoic acid Receptors (ACTR) that interact with the C-terminus of CBP via the CBP_{NCBD}, but not by cAMP Response Element-Binding protein (CREB) that binds to the N-terminal CBP_{KIX} domain[29].

Our data has broader implications for the deregulation of CBP function in disease. Importantly, many variants in CBP associated with diseases such as the developmental disorders Rubinstein Taybi Syndrome (RTS)[67,68] and Menke-Hennekam Syndrome (MHS)[69], and cancers such as B-cell lymphoma[70], map to IDRs characterised here. Disease-associated variants in the HAT domain of CBP are understandably the best characterised, as they are more prevalent and lead to easily definable changes in HAT activity[68,70–72]. However, increasing evidence highlights how variants in IDRs can also lead to disease, including by changing patterns of

605 condensate formation[73]. This study defines clear roles for CBP-IDRs in regulating diverse cellular functions,
606 which if disrupted by mutations could contribute to disease etiology. Indeed, many diseases arising from
607 variants in CBP are heterozygous[70], and therefore the ability of CBP-IDRs to confer dominant
608 positive/negative effects on endogenous CBP condensates by disrupting homotypic interactions may be
609 relevant to these mixed populations. We also highlight significant sequence divergence between functionally
610 relevant IDRs in CBP and their counterparts in p300. We speculate that these differences in sequence could
611 help to explain the contrasting behaviour of CBP_{wt}-GFP (Figure 4E-G) and $p300_{wt}$-GFP condensates following
612 treatment with A-485[11]. Although not tested here, it's possible that these less well-conserved, but functionally
613 important domains, could form the basis for functional specialisation of CBP and p300.

614 Limitations

615 There are a number of limitations to our study. Firstly, as we did not deplete endogenous CBP, we cannot rule
616 out the possibility of complexities arising from mixed populations of mutant and wild-type CBP. Our results
617 show how CBP-IDR mutations can affect the behaviour of endogenous CBP condensates (Figure 2L), but we
618 did not look at how CBP-IDRs affected condensates formed by other factors or address the contribution of
619 heterotypic interactions to CBP condensates; given the complex nature of condensates, which are composed of
620 a plethora of factors including both CBP and p300[10,11], we might expect that larger heterotypic interaction
621 networks are similarly disrupted. This has further implications for changes in histone acetylation, as the impact
622 of CBP-IDR mutations on H3K27ac is likely complicated by the recruitment of endogenous copies of CBP_{wt}
623 and potentially p300 to the same loci.

624 Secondly, our initial studies used transient co-expression of CBP-GFP to assess condensate behaviour. This
625 system is imperfect, as the formation of condensates is highly sensitive to protein concentration, and
626 differences in transfection efficiency could therefore lead to inconsistent behaviour at a cell-to-cell level.
627 Nevertheless, we believe these approaches accentuate behaviours that can also occur at near-endogenous
628 expression levels, as demonstrated using inducible CBP expression (Figure 4L).

629 We cannot rule out the possibility that changes to CBP condensates result from the gain or loss of TF binding,
630 or changes to PTMs on CBP-IDRs. CBP_{IDR6} and CBP_{IDR7} bridge CBP_{NCBD}, a TF binding domain that associates
631 with different TFs[15–19]. It is possible that disrupted CBP_{NCBD} folding, changes TF binding profiles and that this
632 could explain changes in condensate behaviour, although we note that deletion the equivalent region in p300
633 only mildly impacted p300 chromatin association[65].

634 Finally, CBP_{AL} has a variety of roles that could each individually impact CBP condensates, making
635 experiments focussed on CBP_{AL} challenging to interpret. Firstly, CBP_{AL} is a key binding site for RNAs in the
636 nucleus[22]; given the importance of RNA in condensate formation[59], it is likely that disrupting interactions
637 between CBP_{AL} and RNA could alter condensate properties. Indeed, our data suggest that RNA helps to
638 maintain CBP condensates (Figure 2J-K). Secondly, acetylation of lysines in CBP_{AL} directly regulates CBP's
639 catalytic activity[21], potentially altering global lysine acetylation with implications for condensates[56,74].

CBP is fundamentally important for controlling gene expression, integrating diverse environmental signals to generate precise histone acetylation profiles and scaffold assembly of transcription complexes[10,12]. Our work highlights how CBP-IDRs facilitate this process, regulating CBP condensate properties and changing how CBP responds to other environmental factors such as lysine acetylation and changing chromatin localisation, histone acetylation and gene expression patterns. We therefore underline the importance of understanding the unique behaviours of CBP-IDRs and IDRs in other multivalent transcription coactivators when considering their role in gene expression.

Resource availability

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Daniel Bose (d.bose@sheffield.ac.uk). All plasmids generated for this study are available from Addgene. All unique/stable reagents generated in this study are available from the lead contact without restriction. The high-throughput sequencing data data from this publication have been deposited to the GEO database (https://www.ncbi.nlm.nih.gov/geo/) and assigned the identifiers GSE268266 (ChIP-seq), GSE299613 (CUT&RUN) and GSE299614 (RNA-seq). Datasets GSE268266 GSE299613 GSE299614 have been grouped under **SuperSeries GSE299618**.

Acknowledgements

D.B. and AT are Sir Henry Dale Fellows, funded by the Wellcome Trust and the Royal Society (grant numbers 213501/Z/18/Z and 220192/Z/20/Z respectively); K.G., D.B., A.T. and T.D.C. are supported by a BBSRC Pioneer Award (grant number: BB/Y513453/1); K.G. was supported by a studentship from the White Rose BBSRC Doctoral Training Partnership (grant number: BB/M011151/1**)**; L.J.H. was supported by a studentship from the Medical Research Council Discovery Medicine North (DiMeN) Doctoral Training Partnership (grant number: MR/N013840); N.A.C and G.G were supported by UoS PhD studentships and N.A.C by a UoS publication scholarship; T.D.C was funded by (BB/T008032/1); D.B also received funding from The Royal Society (grant number: RSG\R1\180410). M.D. is part of the Sheffield Bioinformatics Core team. At the University of Sheffield, we thank Susan Clark for support at the Flow Cytometry Core Facility; IT services for access to High Performance Computing and the Wolfson Light Microscopy Facility. The Nikon W1 spinning disk confocal was funded by a BBSRC ALERT2021 award (grant number: BB/V019368/1). This research was funded in whole, or in part, by the Wellcome Trust [213501/Z/18/Z and 220192/Z/20/Z]. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Author contributions

The study was conceived and led by D.A.B, A.E.T, T.D.C and K.L.G. D.A.B, A.E.T and K.L.G designed experiments with input from T.D.C and N.A.C. K.L.G carried out all imaging experiments; N.A.C generated

tagged cell lines and prepared samples for high-throughput sequencing data; L.J.H, M.D and D.A.B analyzed NGS data; K.L.G, N.A.C, G.G, L.J.H, B.V.E, T.E, S.B carried out all other experiments. D.A.B., K.L.G and A.E.T wrote the manuscript. All authors reviewed and commented on the manuscript.

Disclosure and competing interests statement

Timothy Craggs is the founder and CEO of Exciting Instruments, a company that develops and sells instrumentation for single-molecule fluorescence experiments, including smFRET.

Supplemental information

Supplemental information includes: 7 supplemental figures, 4 tables, 1 document containing pairwise sequence alignments, 8 movies.

Methods

Materials and Methods

Cell culture.

Culture and maintenance.

Human Embryonic Kidney cells stably expressing the SV40 large T antigen (HEK293T, ATCC) and GP2-293 (Clontech) cells were cultured in complete DMEM (10% FBS, 1% Penicillin-Streptomycin (Gibco)) and grown in 5% CO_2 at 37ºC. HEK293T cells containing stable doxycycline inducible CBP-GFP, were cultured in complete DMEM (10% Tetracycline free FBS, 1% Penicillin-Streptomycin) including 2 mg/ml G418 (Generon) and 50ug/ml hygromycin (ThermoFisher). All cell lines were screened every three months for mycoplasma infection.

Transient transfections.

Transient transfections were done using either Lipofectamine 3000 (ThermoFisher Scientific) or LipoD293 (TEBU-Bio). Manufacturers' recommended protocol was followed for 6-well dishes with the following details: 1.5 µg of DNA was transfected per well and 7.5 µl of Lipofectamine 3000 reagent was used per well. For LipoD293 (TEBU-Bio), the manufacturers protocol for 6-well 35-mm dishes was followed using: 1 µg of DNA and 3 µl of LipoD293 per well.

Retroviral transduction.

Lentiviral transduction of CBP-GFP was done using the pRetroX^{TM} Tet-On^{R} 3G inducible expression system (Clontech). CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP, and $CBP_{\Delta IDR7}$-GFP were cloned into the pRetroX-pTRE3G vector for lentiviral transduction. To generate lentiviral particles, 2 µg of CBP-pRetroX-pTRE3G constructs or pRetroX-Tet3G were co-transfected with 2 µg of the envelope pVSV-G plasmid into GP2-293 cells using lipofectamine 3000 (ThermoFisher). After 48 hours, with 1 media change after 24 hours, the viral supernatant was collected and concentrated using Lenti-X concentrator (Takara) following manufacturers instructions. To generate polyclonal pRetroX-Tet3G HEK293T cells, 8 ug/ml polybrene (Merck Millipore) was added to the concentrated virus and used to infect HEK293T cells. Infected cells were selected using G418 (Generon) at a concentration of 8 mg/ml for 1 week. To generate CBP-pRetroX-pTRE3G cells, polyclonal pRetroX-Tet3G HEK293T cells were transduced with viral particles and 8 ug/ml polybrene, then selected using 8 mg/ml of G418 and 125 µg/ml hygromycin (ThermoFisher) for 1 week. Clonal cell lines were generated by inducing

expression of CBP-GFP with 2000 ng/ml of doxycycline for 48 hours before fluorescence-activated cell sorting (FACS) based on the GFP signal, using a FACSMelody Cell Sorter (BD Biosciences) at the University of Sheffield Flow Cytometry Core Facility. Clonal lines were cultured in the background of 2 mg/ml G418 and 50 µg/ml hygromycin (ThermoFisher).

Doxycycline induction of CBP expression.

For doxycycline validation experiments, HEK293T cells were plated in 10 cm dishes (Sarstedt) for western blot or 35-mm glass bottom dishes (Ibidi) for imaging. 24 hours post-plating the media was changed to media containing the predetermined concentration of doxycycline calculated to ensure induction of CBP-GFP but not increase overall CBP expression levels (Figure S4K-L)), 2 mg/ml G418 and 50 µg/ml hygromycin.

Molecular Biology and cloning.

All PCR reactions were completed using a 50 µl reaction containing: 0.5 µl PFuUltra II fusion HS DNA Polymerase (Agilent), 1 % PFu Ultra reaction buffer, 5 % dimethyl sulfoxide (DMSO, Sigma), 200 mM dNTP mix, 0.2 µM forward and reverse primer, and template DNA. DNA was initially denatured at 95ºC with an extension temperature of 68ºC for DNA less than 10 Kb, and 72ºC for DNA greater than 10 Kb, with 35 extension cycles unless otherwise stated. To confirm correct assembly and presence of mutations, all plasmids were checked by Sanger sequencing and most were checked by Whole plasmid sequencing (Eurofins). All Constructs have been made available through Addgene (https://www.addgene.org/).

Full length human CBP (CBP_{wt}).

To generate human CBP_{wt} (transcript variant 1, NCBI Reference: NM_004380.3) tagged with a C-terminal GFP for overexpression experiments (CBP_{wt}-GFP), *Homo sapiens* CBP transcript variant 2 (NCBI Reference NM_001079846.1, Sino Biological HG17295-UT) was cloned in-frame into pAcGFP-N1 (Clontech) using NEBuilder HiFi DNA Assembly reaction protocol (NEB). CBP was amplified using primers: *Fwd: Human CBP_fwd, Rev: Human CBP_rev* and pAcGFP-N1 was amplified using primers: *Fwd: pAcGFP-N1_fwd, Rev: pAcGFP-N1_rev*; (For primer sequences see Supplementary table 1). Assembled constructs were checked by Sanger sequencing (Eurofins). To convert transcript variant 2 to transcript variant 1, the missing sequence comprising exon 5 was ordered as a custom-synthesised DNA from IDT and inserted into CBP-pAcGFP-N1 using the NEB HiFi assembly. CBP exon 5 was amplified using primers: *CBP missing Exon_fwd, CBP missing Exon_rev*. CBP-pAcGFP-N1 was amplified using primers: *Fwd: GFP-CBP_fwd, Rev: GFP-CBP_rev*. Fully assembled CBP_{wt}-GFP was sequenced by Sanger sequencing using sequential primers at 900bp intervals and confirmed using Whole Plasmid Sequencing (Eurofins) to check for correct assembly.

CBP-IDRs for optoDroplet.

The optoDroplet construct pHR-mCh-Cry2WT was a gift from Clifford Brangwynne (Addgene plasmid #101221)[39]. To target the optoDroplet construct to the nucleus we used custom synthesised DNA (gBlock, IDT) encoding the SV40 nuclear localisation signal (NLS) to generate pHR-mCh-Cry2WT-NLS [36]. The SV40 NLS was amplified by PCR using primers: *Fwd: SV40_NLS_NotIRD_fwd, Rev: SV40_NLS_SbfIRD_rev*, to introduce restriction sites for NotI and SbfI. Following restriction digestion with NotI-HF and SbfI (NEB), pHR-mCh-Cry2WT and the synthesised NLS were ligated using T4 DNA ligase (NEB) and the construct was checked by Sanger sequencing. CBP-IDRs were cloned into pHR-mCh-Cry2WT-NLS construct using the NEB HiFi DNA assembly protocol. pHR-mCh-Cry2WT-NLS was amplified using primers: *Fwd: pHR-mCh-Cry2WT-NLS_fwd, Rev: pHR-mCh-Cry2WT-NLS_rev*. CBP-IDRs were amplified from CBP_{wt}-GFP using IDR-specific primers (Supplementary table 1). All constructs were checked using Sanger sequencing.

$CBP_{\Delta IDR}$-GFP deletion mutants.

To generate $CBP_{\Delta IDR}$ mutants, primers were designed to amplify CBP_{wt}-GFP in reverse directions to delete residues comprising the identified CBP-IDRs (Supplementary table 1). Amplified PCR products were digested using DpnI (NEB), the ends of the PCR fragments were then phosphorylated using T4 polynucleotide kinase (PNK, NEB) before undergoing ligation using T4 DNA ligase (NEB). Constructs were sequenced by Sanger sequencing to check the IDRs had been correctly removed from the construct.

Autoregulatory Loop (CBP_{AL}) mutants.

Two mutants of the Autoinhibitory Loop (CBP_{AL}, $CBP_{1558-1607}$) were designed. Firstly, CBP_{KTG} contained the AL with all lysines mutated to glycine (K1564G, K1565G, K1583G, K1586G, K1587G, K1588G, K1591G, K1592G, K1595G, K1597G, K1605G, K1606G, and K1607G). Secondly, in $CBP_{\Delta AL}$ the AL was replaced with a glycine/serine linker (sequence GSAGSAAGSGQF) to maintain correct folding[23]. $CBP_{4322-5818}$ - which are flanked by unique restriction sites for NruI and PmlI - were synthesised as gBlocks (IDT) containing the mutant AL sequences (Supplementary table 1). The synthesised gBlocks and CBP-pAcGFP-N1 constructs (CBP-GFP for single mutants; $CBP_{\Delta IDR6}$-GFP, $CBP_{\Delta IDR7}$-GFP, and $CBP_{\Delta CFID}$-GFP for double mutants) were digested using NruI-HF and PmlI (NEB). The digested constructs were dephosphorylated using Quick Calf Intestinal Phosphatase (CIP, NEB), before the inserts and vectors were assembled using T4 DNA ligase (NEB). The assembled constructs were checked by Sanger sequencing.

Doxycycline inducible CBP-GFP.

CBP-GFP, $CBP_{\Delta IDR6}$-GFP, and $CBP_{\Delta IDR7}$-GFP, were sub-cloned into a modified pRetroX-pTRE3G plasmid (Clontech) containing a hygromycin resistance cassette using BamHI-HF and NotI-HF (NEB) restriction enzymes. The fragments were annealed using T4 DNA ligase (NEB) before sequencing to check for correct insertion and subsequently by whole plasmid sequencing.

782 $CBP_{\Delta IDR6H2}$-GFP and Cloning $CBP_{\Delta IDR7H1}$-GFP.

783 To delete predicted a-helices in CBP_{IDR6} and CBP_{IDR7}, DNA sequences were synthesised containing CBP 5332
784 - 7674 which resides within the N-terminal GFP tag, and spans two unique endogenous restriction sites for
785 MluI-HF and BssHII (GeneArt, ThermoFisher). Synthesized sequences contained deletions of $CBP_{1951-1978}$
786 ($CBP_{\Delta IDR6H2}$-GFP) and $CBP_{2187-2216}$ ($CBP_{\Delta IDR7H1}$-GFP). CBP_{wt}-GFP and synthesised sequences were digested
787 with MluI-HF and BssHII (NEB) and purified by agarose gel extraction. The digested CBP_{wt}-GFP vector was
788 dephosphorylated using Quick Calf Intestinal Phosphatase (CIP, NEB) before assembly using T4 DNA ligase
789 (NEB). Correct insertion was checked by Sanger sequencing.
790

791 $CBP_{IDR6\text{-}Shuffle}$-GFP and Cloning $CBP_{\Delta IDR7\text{-}Shuffle}$-GFP.

792 DNA sequences were synthesised of CBP_{IDR6} and CBP_{IDR7} where the overall sequence composition was
793 maintained but the sequence was randomly shuffled to disrupt the sequence patterning. Design of the shuffled
794 sequences was performed using GOOSE [52] to randomize the input sequence amino acid sequence. DNA
795 coding for the shuffled sequences was designed and codon optimised for expression in *H. sapiens* (GeneArt,
796 ThermoFisher). The synthesised DNA contained two unique restriction sites for MluI-HF and BssHII, cloning
797 was performed as described for $CBP_{\Delta IDR6H2}$-GFP and Cloning $CBP_{\Delta IDR7H1}$-GFP.

798 Cloning of gRNA plasmids.

799 Designed gRNAs (Supplementary table 1) were cloned into pSpCas9(BB)-2A-GFP (PX458). PX458 was a
800 gift from Feng Zhang (Addgene plasmid # 48138) [75]. 100 µM forward and reverse oligo pairs were combined
801 with T4 polynucleotide kinase (PNK) (NEB) and 10X T4 ligation buffer (NEB) containing essential ATP. For
802 annealing, reactions were heated to 37°C for 30 minutes, followed by 95°C for a further 5 minutes and then
803 cooled to 25°C at a ramp rate of -5°C per minute. Annealed gRNAs were then diluted 1:100 in sterile H_2O in
804 preparation for the ligation reaction.
805 pX458 plasmid was digested with BbsI-HF (NEB) DNA and annealed gRNA oligos were ligated using T4
806 DNA ligase (NEB). Constructs were validated using Sanger sequencing.

807 Construction of stably integrated CBP-Halotag cell lines.

808 gRNA design.

809 Single guide RNAs (sgRNAs) targeting the C-terminus of CBP were designed using the UCSC CRISPR
810 targets tool (https://genome.ucsc.edu/). After selection in UCSC, crRNAs were further evaluated using the
811 Evaluation tool on the E-CRISP site (http://www.e-crisp.org/E-CRISP/reannotate_crispr. html). Plasmids
812 encoding CBP-targeting sgRNAs were transfected into HEK293T cells using the Neon electroporation system
813 and 100 µl kit (ThermoFisher), at 1100V, 20 ms pulse width and 2 pulses. After 48 hours, genomic DNA was
814 extracted using QuickExtract solution (Lucigen) according to the manufacturer's instructions and PCR of CBP
815 C-terminal region was carried out using primers *CBP_C-T_fwd1/2, CBP_C-T_rvs* (Supplementary Table 1).

816

817 gRNA validation.

818 PCR products amplifying the C-terminal region of CBP in both edited and unedited cells were cleaned up
819 using the Monarch PCR cleanup kit (NEB). DNA concentrations were measured using the Qubit dsDNA BR
820 kit (ThermoFisher) according to the manufacturer's instructions. 10X NEB buffer 2 (NEB) was added to 200
821 ng DNA and PCR products were denatured to single strands by heating at 95ºC for 5 minutes. DNA strands
822 were then hybridised by decreasing temperature from 95ºC to 85ºC at a ramp rate of -2ºC/second, then from
823 85ºC to 25ºC at a ramp rate of -0.1ºC per second. T7 Endonuclease 1 (T7E1) enzyme (NEB) was added and
824 incubated at 37ºC for 15 minutes, followed by quenching with 0.25 M EDTA. The total reaction volume was
825 run on a 1% agarose gel. To analyse Sanger sequencing data of CRISPR edited DNA sequences, .ab1 files of
826 CBP C-terminal PCR products were uploaded to the Inference of CRISPR Edits (ICE, Synthego) tool
827 (https://ice.synthego.com/#/, Conant et al. 2022).
828

829 Generation of clonal CBP-Halo HEK293T cell lines.

830 Validated plasmids encoding CBP gRNA and custom synthesised CBP HDR template plasmid (Supplementary
831 Table 1, ThermoFisher) were co-transfected into HEK293T cells using the Neon electroporation system and
832 100 µl kit (ThermoFisher) at 1100V, 20 ms pulse width and 2 pulses. After 48 hours, single cell sorting was
833 performed using a FACSMelody cell sorter (BD Biosciences) into 96 well plates. Clones were allowed to grow
834 until confluent and then screened PCR genotyping for successful integration of the Halotag using primers
835 *CBP_C-T_fwd1/2, CBP_C-T_rvs* (Supplementary Table 1). We performed sufficient screening to identify
836 heterozygous clones. However, we cannot eliminate the possibility that the C-terminal HaloTag caused
837 detrimental effects for CBP function that resulted in no homozygotes being obtained.
838

839 qPCR.

840 Total RNA was extracted from WT and CBP-HaloTag HEK293T cells, using TRIzol (ThermoFisher) and
841 treated with Turbo DNase (Invitrogen) to remove genomic DNA, then purified using phenol/chloroform and
842 ethanol precipitation. Reverse Transcription was performed using the High Capacity cDNA kit
843 (ThermoFisher) and qPCR was carried out using Power SYBR reaction mastermix (ThermoFisher) and a
844 Quantstudio 12K Flex (ThermoFisher). A mean Ct was calculated using each triplicate reaction and
845 normalised to 18S rRNA expression.
846

Imaging

Live cell imaging of CBP-HaloTag.

HEK293T cells containing endogenous CBP-HaloTag were seeded at 400,000 cells/well on 35 mm glass bottom dishes that had been coated in 20 µg/ml of poly-L lysine (PLL, Merck) diluted in sterile dH_2O for 1 hour at 37ºC, before washing twice in sterile dH_2O. After 48 hours, cells were labelled with HaloTag ligand TMR (Promega) at a concentration of 50 nM for 15 minutes at 37ºC in the dark. Cells were left for a further 2 hours before staining with Hoechst 33342 (Fisher) in DPBS as described in a previous section, before washing in DPBS and storing in imaging media (FluoroBright DMEM (Fisher) supplemented with a final concentration of 20 nM Gibco HEPES (ThermoFisher Scientific)) immediately prior to imaging. Cells were imaged using the Nikon W1 spinning disc confocal microscope equipped with a temperature stage set to 37ºC. Images were taken using the 100X oil immersion lens (NA 1.45), using the 405 nm laser to image Hoechst using 45 % laser power with a 100 ms exposure, TMR was visualised using the 514 nm laser using 60 % laser power with an exposure of 300 ms. Images were taken every 500 ms in the absence of binning. Captured images and videos were processed ImageJ/Fiji[76].

FRAP of CBP-HaloTag.

CBP-HaloTag cells were plated on PLL coated 35-mm dishes, and labelled with Janelia Fluor® 549 (JF549, Promega) at a concentration of 50 nM for 15 minutes. Cells were subsequently washed twice in DPBS and placed in imaging media for a 15 minute recovery. Cells were imaged using a custom built iLas TIRF/FRAP single-molecule scanning microscope by Cairn. Puncta were imaged using 561 nm laser with 200 ms exposure at 10 % laser power, on the 100 % oil objective. ROIs were drawn with a diameter of 10 to mark where the lasers should target, this could be multiple puncta per field of view. FRAP settings are as follows: 10 repetitions using the 561 nm laser line of weight 2 on 3 % laser power. A time course was taken with the following pattern: 15 time points with a 500 ms delay prior to FRAP which was performed as described above, followed by 60 time points with a 500 ms delay and 30 time points with a 2 second delay. Images were quantified and processed in ImageJ/Fiji [76].

Imaging of transfected CBP-HaloTag with CBP-GFP, $CBP_{\Delta IDR6}$-GFP, $CBP_{\Delta IDR7}$-GFP

Endogenous CBP-HaloTag cells were seeded at 250,000 cells/well on PLL coated glass bottom 35-mm dishes and transfected after 24 hours with CBP-GFP, $CBP_{\Delta IDR6}$-GFP, $CBP_{\Delta IDR7}$-GFP using LipoD293, as has been previously described. Cells were labelled with Janelia Fluor® 646 (JF646, Promega) at a final concentration of 2 nM for 15 minutes before washing in DPBS and imaging in imaging media. 35 mm dishes were labelled sequentially, immediately before imaging.

Imaging using custom built iLas TIRF/FRAP single-molecule scanning microscope by Cairn.

4-colour control images were also taken where the HaloTag had not been labelled, to check for bleed-through of overexpressed constructs. In this case, the media was changed to imaging media and imaged using laser lines 405 nm, 488 nm, 532 nm and 637 nm, with an exposure time of 200 ms using 10 % laser power. A 4-colour control was also taken with CBP-HaloTag labelled with JF646 and no overexpression, for this all laser lines used the same conditions as before except the 637 nm laser had a power of 80 % to visualise the endogenous protein. For colocalization experiments, 2 colour multidimensional images were captured using the 488 nm laser with 10 % laser power and a 200 ms exposure to visualise GFP tagged constructs, and 637 nm laser with 80 % laser power and 200 ms exposure to view endogenously tagged HaloTag. Images were processed in ImageJ/Fiji [76].

Imaging using Zeiss Lattice Lightsheet 7.

3D images of CBP-HaloTag and overexpressed CBP-GFP, $CBP_{\Delta IDR6}$-GFP, $CBP_{\Delta IDR7}$-GFP were taken using the Zeiss Lattice Lightsheet 7, using the Sinc3 30 x 1000 lightsheet with a calibrated magnification of 50. The system was calibrated following manufacturers guidelines. 488 nm laser line was used to image overexpressed GFP tagged constructs at 40% laser power using 50 ms exposure, the endogenous CBP-HaloTag was imaged using the 638 nm laser line using 80 % laser power with a 100 ms exposure. Zen Blue software was used to deskew the dataset and for representative images channel alignment was performed using default settings, where the CBP-HaloTag (638 nm images) were used for reference. Images were then pre-processed in ImageJ/Fiji [76], before quantification using Arivis Vision4D version 3.4.

Imaging of doxycycline inducible clones.

For imaging cells were seeded at a concentration of 250,00 cells/well on PLL coated 35-mm dishes and doxycycline induced for 48-hours. Imaging was performed in imaging media using the Nikon W1 spinning disc confocal microscope, using the 100X oil objective. Laser lines 405 nm and 488 nm were used to image the Hoechst 33342 (Fisher) and therefore the nuclei, and GFP expression respectively. The 405 nm laser was set with a 100 ms exposure with 60 % laser power, and the 488 nm laser was set with 300 ms exposure time and 60 % laser power. For quantification Z-stacks were taken to capture the whole volume of the nuclei in 0.5 µm increments, these were further processed in ImageJ/Fiji [76]. For the plus doxycycline conditions, imaging media was supplemented with the appropriate concentration of doxycycline.

optoDroplet.

35-mm glass bottomed dishes (Ibidi) were coated in PLL for 1 hour at 37ºC before being washed twice with dH_2O and seeded with HEK293T cells at a concentration of 200,000 cells/well. 24 hours post-plating the cells were transfected using LipoD293™(SignaGen Laboratories), following manufacturer's guidelines for 35 mm dishes. After 24 hours the media was changed into complete DMEM (10% FBS, 1% Pen/Strep). After a further 24 hours, prior to imaging the media was changed into FluoroBright DMEM (Fisher) with a final concentration of 20 mM HEPES (ThermoFisher).

914 For IDR1, IDR3, IDR4, AL, IDR6, IDR7 and the CFID the live cell imaging was performed on Nikon A1
915 confocal using the 60X oil immersion objective (NA 1.4), equipped with a humidified temperature stage set to
916 37ºC. For the IDR2 and IDR5 live cell imaging was performed on the Nikon W1 spinning disc confocal
917 microscope, using a temperature stage set to 37ºC and the 100X oil objective lens. In both cases, two laser
918 wavelengths were used to induce global activation of the optoDroplet system; 488 nm to induce the
919 dimerisation of the Cry2 and 560 nm to image mCherry. An image was captured for the mCherry every 8
920 seconds, with subsequent activation of the mCherry for the same time. Time courses were taken for each
921 construct, for a period of up to 3 minutes. 3 time courses were taken per construct per biological replicate, with
922 a total of 3 biological replicates being taken.
923

924 Fluorescence recovery after photobleaching (FRAP) in transfected cells.

925 For FRAP experiments, 250,000 cells/well of HEK293T cells were plated in 35 mm dishes and transfected
926 using LipoD293 as was described for the optoDroplet live cell imaging. 24-hours post transfection the media
927 was removed and the cells were washed in DPBS before leaving the cells in imaging media. Imaging was
928 performed using a custom built iLas TIRF/FRAP single-molecule scanning microscope by Cairn, using an
929 100X oil immersion objective. Puncta was visualised using the 488 nm laser set to 5 % laser power with a 100
930 ms exposure. A circular ROI with dimensions of 4.4 x 4.4 µm was drawn over the puncta to be
931 photobleached; there may be multiple regions within each field of view. Photobleaching was performed using
932 the 488 nm laser, with 10 % laser power, using 5 repetitions with a thickness of 5. 10 images were taken prior
933 to FRAP, followed by 700 frames being collected after, with a frame interval of 0.387 seconds.

934 1,6 - Hexanediol treatment.

935 Cells were plated and transfected as previously described for FRAP in transfected cells. Images were captured
936 using the Nikon W1 spinning disc confocal microscope using the 100X oil immersion objective. Images were
937 taken using the 488 nm laser with a 200 ms exposure time on 20 % laser power. A 10% solution of 1,6 -
938 Hexanediol (Merck) was diluted from a 50% stock in warm imaging media. Time courses were taken of the
939 constructs over a 3 minute period, imaging every 500 ms. Without pausing imaging, after 30 seconds 1 ml of
940 10% 1,6 - Hexanediol was added dropwise to the 35 mm dish that contained 1 ml of imaging media for a final
941 concentration of 5% 1,6 - Hexanediol. Imaging continued until the end of the time course. Image
942 quantification and processing was performed in ImageJ/Fiji [76].
943

944 RNAse treatment.

945 Cells were plated and transfected in 35mm glass bottom plates as described above. On the day of imaging, the
946 media was removed and replaced with FluoroBrite media and images were collected of the untreated cells. The
947 cells were then RNAse treated[77] as follows: firstly the cells were washed with PBS and reaction buffer (20mM

Tris pH 7.5, 5 mM $MgCl_2$, 0.5mM EGTA, 1x protease inhibitor cocktail) before being permeated with 0.1% Triton X-100 for 5 mins. A second round of washes was then applied before 20 minute incubation with 0.2 mg/ml final RNase A (NEB), diluted in the nuclease buffer (5 mM $MgCl_2$ in PBS) to a final concentration of 100 µg/ml. After treatment the RNase solution was removed and replaced with imaging media. Images were subsequently collected of the treated cells. Imaging was performed using the Nikon W1 spinning disc confocal microscope, using the 100X oil objective. The 488 nm laser line was used to image GFP expression. For quantification Z-stacks were taken to capture the whole volume of the nuclei in 0.5 µm increments, these were further processed in ImageJ/Fiji [76].

Immunofluorescence and fixed cell imaging.

Immunofluorescence and fixed cell imaging.

Circular coverslips (Ibidi) were acid washed in 0.25% acetic acid before washing and storing in 100% ethanol. Coverslips were coated in poly-L lysine (PLL, Merck) for 1 hour at 37°C before washing twice in sterile dH_2O. HEK293T cells were grown in 6-well plates at a concentration of 200,000 cells/well containing two coverslips per well for 48 hours before fixing. If the cells were transfected, transfection was performed 24 hours after plating using Lipofectamine 3000 transfection reagent (ThermoFisher) or LipoD293 (TEBU-Bio) following manufacturers instructions. Cells were washed in DPBS twice before fixing in 2% paraformaldehyde (PFA) for 30 minutes at room temperature in the dark, before being washed 3 times in DPBS. For immunofluorescence, coverslips were blocked using a blocking buffer (5% Bovine serum albumin (BSA) with 0.2% Triton in PBS) for 1 hour at 37°C. Primary antibodies (for full list of antibodies see Supplementary Table 2) were prepared in a blocking buffer and incubated at 37°C for 1 hour. After washing 3 times in PBS, secondary antibodies were diluted in the blocking buffer and incubated for 37°C for 1 hour. After fixing, coverslips were mounted onto slides (VWR, 631-1554) using VECTORshield that contained 4,6-diamidino-2-phenylindole (DAPI) (Vector laboratories, H-2000-10). Fixed slides were stored at 4°C. Imaging work was performed at the Wolfson Light Microscopy Facility using the Nikon A1 confocal microscope. All fixed cell imaging was performed using the 60X oil immersion objective (NA 1.4). The microscope has 4 laser wavelengths, typically only 3 of the 4 were used (these include, 405 nm for DAPI, 488 nm for GFP, 562 nm for mCherry and 642 nm for Alexa Fluor 647). Z-stacks were taken every 0.5 µm to cover the depth of the nuclei; a minimum of 3 fields of view were imaged per construct over 2 coverslips, per biological replicate.

For CBP deletion mutants $CBP_{\Delta IDR2}$ and $CBP_{\Delta IDR5}$ images were captured on the Nikon W1 spinning disc confocal microscope, using the 100X oil objective and the 488 nm laser line. Z-stacks were taken as described above.

A-485 treatment

Cells were cultured in 6-well plates, containing PLL-treated circular coverslips as described above. HEK293T cells were transfected with LipoD293 (TEBU-Bio) after 24 hours following manufacturers instructions. For 36 hour A-485 treatment, media was removed after 8-hours of transfection and replaced with media supplemented with A-485 to a final concentration of 5 uM, or with media containing DMSO as a negative control. The cells were treated for 36 hours before fixing with 2 % paraformaldehyde and mounted on slides using VECTORshield containing DAPI as described above.

For the A-485 concentration experiments, cells were transfected after 24 hours of culture, and treated 24 hours after transfection. Media was replaced either 2 hours before fixing or 1 hour before fixing with media containing A-485 at concentrations of 1 uM, 5 uM or 10 uM. A DMSO control was performed at each time point. After treatment cells were fixed as described above. Imaging was performed using the Nikon W1 spinning disc confocal microscope, using the 100X oil objective. Laser lines 405 nm and 488 nm were used to image the DAPI and therefore the nuclei, and GFP expression respectively. For quantification Z-stacks were taken to capture the whole volume of the nuclei in 0.5 µm increments, these were further processed in ImageJ/Fiji [76].

TSA treatment

Cells were cultured in 6-well plates, containing PLL-treated circular coverslips as described above. HEK293T cells were transfected with LipoD293 (TEBU-Bio) after 24 hours following manufacturers instructions. After ~36 hours cells were treated with Trichostatin A (TSA, Stratech) at a final concentration of 500 nM for 90 minutes. After 90 minutes the cells were fixed and imaged as described above for A-485 treatment.

Image analysis

Integrated Intensity Density

Before calculating the Integrated density, individual nuclei were isolated from a maximum projection and saved as .tif files where the file names were blinded using ImageJ/Fiji plugin: Blind Analysis Tool, File Name Encryptor [76]. These blinded files were categorized by two individuals independently, scoring the images as either diffuse, containing puncta, overexpressed or untransfected. Any nucleus that had a single puncta , was described as containing puncta; those that were transfected but displayed a diffuse signal were categorized as diffuse. Overexpressed nuclei were any nuclei where the signal intensity was too great to identify individual puncta, and untransfected cells were nuclei where little to no signal intensity was observed (Figure S2B). Both overexpressed and untransfected nuclei were removed from the dataset and an average percentage of nuclei which displayed either puncta or a diffuse phenotype was determined for each of the constructs.

To assess the punctate nature of the nuclei in an unbiased manner we also calculated the standard deviation of pixel intensity across the nucleus. Nuclei containing puncta would display a higher standard deviation between

pixels compared to nuclei with a diffuse signal. A region of interest (ROI) was drawn around the nucleus by applying a manual threshold to isolate only the nuclei in ImageJ/Fiji [76]. This region was applied to the raw image and the standard deviation was calculated within this region. All downstream analysis was done using R in Rstudio; significance p values for comparison of punctate signals were calculated using a Kruskal-Wallis test. Graphs and statistical testing was performed using R in RStudio [78].

Integrated density, the intensity multiplied by the area, for each condensate produced by a defined construct was calculated in ImageJ/Fiji [76] using a custom macro. A maximum projection was generated from each Z-stack, and the condensates were identified by applying a manual intensity threshold to each image to generate a mask. The intensity threshold was internally consistent within each experimental repeat, where the threshold was chosen to best isolate condensates across the range of different phenotypes observed within the constructs tested. Within the mask, each independent ROI was labelled, and the area, Integrated Density, and circularity was measured for each ROI.

Once these images had been analysed, the percentage of nuclei which contained a ROI which has an intensity higher than that of the set intensity threshold value was further calculated. This value was described as the transfection efficiency, and was generated to identify which nuclei had passed the threshold value. Filters were then applied sequentially to the data: 1) To exclude overexpressed nuclei, a maximum grey value of 10 was set; 2) To remove regions of high intensity that were too small to represent puncta, an area threshold (<3 pixels/0.36um) was applied; 3) To remove regions of overlapping puncta, a circularity threshold (<0.25) was set. All downstream analysis was done using R in Rstudio; significance p values for comparison of CBP-GFP behaviours (number of condensate per nuclei and integrated density) were calculated using a Kruskal-Wallis test. Graphs and statistical testing was performed using R in RStudio [78].

optoDroplet

A custom ImageJ macro was written to identify puncta formed in the mCherry channel. Images were run through an ImageJ plugin called Stack Registration (StackReg, [79]), and a threshold was applied to generate a binary mask that underwent binary processing of closing followed by opening. For IDR1, IDR3, IDR4, AL, IDR6, IDR7 and the CFID the threshold was set using RenyiEntropy, and for IDR2 and IDR5 the Default setting was used. The parameters were then set in analyse particles, applying a size range of 50 - infinity, to generate ROIs which equated to the outline of the signal region, which was the nuclei. These ROIs were then applied to the original, raw image that underwent a multi-measure to calculate the standard deviation of the pixels within the nuclei over the time course. Regions found in the first frame were used to monitor the change in the standard deviation over the time course, where the average Fold Change was then calculated. A rate of change was then calculated over the time course based on the time at which the curve for the positive control plateau. For IDR1, IDR3, IDR4, AL, IDR6, IDR7 and the CFID this was done over the first 32 seconds, whereas for IDR2 and IDR5 was calculated over the first 72 seconds. For IDR1, IDR3, IDR4, AL, IDR6, IDR7 and the CFID the statistics for the rate of change were performed in GraphPad prism, where an unpaired

t-test was used to determine statistical significance (Prism [80]). For the IDR2 and IDR5 statistics were performed using R in RStudio, using the Kruskal-Wallis test. Graphs for visualisation were made using R in RStudio [78].

FRAP in CBP-HaloTag

Images for each FRAP time course were loaded into ImageJ/Fiji [76] and processed using StackReg ([79]) to remove drift in the sample. Over each time course, the intensity was calculated for a region of interest around the bleached area, before being corrected for photobleaching over the time course using a reference region within the same nuclei. Relative intensity was calculated by subtracting the intensity at T = 0, where this was defined as the first frame after photobleaching, and then normalised to the pre-bleach intensity for each construct. Rate constants and T = ½ were calculated in GraphPad Prism 9 [80], using a non-linear regression with settings one phase association with a variance weighting of $1/Y^2$.

FRAP in transfected cells

Time courses from FRAP of overexpressed protein were quantified using ImageJ/Fiji [76].The intensity was calculated over the time course, using a ROI of 0.22 µm, which is smaller than the condensate to only monitor the intensity within the condensate. The condensate was manually tracked throughout the time course, moving the ROI when the condensate moved. The relative fluorescence recovery was calculated by subtracting the intensity at T = 0, the first frame post bleaching and then normalised to the pre-bleach intensity for each construct. Rate constants and T = ½ were calculated as described for FRAP of CBP-HaloTag.

Quantification of CBP-HaloTag puncta data from the Zeiss Lattice Lightsheet 7 in the absence and presence of overexpressed CBP-GFP, $CBP_{\Delta IDR6}$-GFP, $CBP_{\Delta IDR7}$-GFP

Images were preprocessed in ImageJ/Fiji [76], where background was subtracted using a rolling ball of 50, and the images were sharpened. The preprocessed images were then converted to .sis files using Arivis SIS converter software before being quantified using Arivis Vision4D version 3.4. The data was first imported into the pipeline as a current time point. A denoise preset was run on either the Janelia Fluor channel for CBP-HaloTag only, or on the GFP channel in dual colour images to isolate either all nuclei present or all transfected nuclei respectively. A discrete gaussian of 5 µm was applied to this channel, to blur the edges of the nuclei, after the denoise the results were saved as a temporary document. Next an intensity threshold segmentor was applied to the denoised channel to isolate the nuclei of interest. For CBP-HaloTag only a threshold of 350 was applied, with a core filter of 400 and size range of 500 - 10,000 μm^2 was applied to isolate only nuclei. For overexpressed GFP constructs a threshold of 600 was applied, with a core filter of 1000 and size range of 500 - 10,000 μm^2 was applied. The blob finder was then applied to the Janelia Fluor channel for endogenous CBP-HaloTag in all cases. For CBP-HaloTag only a diameter of 5 µm was applied with a probability threshold of 22 % and a split sensitivity of 50 %. In the dual colour images a diameter of 5 µm was applied with a probability threshold of 6 % and a split sensitivity of 65 % was applied. A segment

feature filter was then applied to remove any objects that are too small to be classified as puncta. A compartment section was then applied to assign the condensates identified in the blob finder to the nuclei identified within the intensity threshold segmentor. The compartmentalisation was selected as a full overlap so that only puncta contained within the nucleus was counted. The data was then exported as a Master-Details report, containing features of the nuclei and details of the condensates - these features included volume, area and intensity of puncta. Graphs for visualisation were made using R in RStudio [78].

Western blotting

For western blotting HEK293T cells were plated in 10 cm dishes (Sarstedt), and induced or treated as described in the imaging section. Cells were harvested and after washing twice in PBS the pellet was lysed in RIPA buffer (50 mM Tris-HCl pH 8 stored at 4ºC, 100 mM NaCl, 2 mM $MgCl_2$, 1 % Triton X-100, 0.1 % Sodium deoxycholate, 0.1 % Sodium Dodecyl Sulfate (SDS)) supplemented with 1 mM Dithiothreitol (DTT), 1X Halt protease inhibitor cocktail, 10 mM sodium butyrate and 500 U/µl benzonase. 25 - 50 µg of overall protein containing sample buffer and reducing buffer was loaded onto either a 3 - 8% Tris-Acetate gel (Thermo Fisher Scientific) using Tris Acetate running buffer (Life Technologies) for large proteins such as CBP, or 4 - 12% Bis-Tris gel with MOPS running buffer for visualising small things such as H3K27ac. Gel was transferred onto a membrane using trans-blot turbo mini 0.2 µm nitrocellulose transfer packs (Biorad) before blocking and probing for CBP, GFP, acetylated CBP 1535, H3K27ac and GAPDH (for full list of antibodies see Supplementary Table 2). Western blots were imaged using G:box (Syngene) gel imager. Images were processed using ImageJ/Fiji [76] and assembled using Adobe Illustrator.

Histone acetyltransferase assay

Expression of CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP, $CBP_{\Delta IDR7}$-GFP was induced in the respective cell lines by 48 hours treatment with previously determined doxycycline concentrations to ensure equal expression levels (Figure 4M & S4K-L). Cells were harvested immediately after doxycycline treatment for cell lysis. Cells were washed with PBS and resuspended in NET buffer (50 mM Tris-HCl, 150 mM NaCl, 2 mM MgCl2, 2 mM CaCl2, 2 mM ZnCl2) supplemented with 0.1% NP40, 10mM sodium butyrate and Halt™ Protease and Phosphatase Inhibitor Cocktail. Lysed cells were incubated with benzonase (1kU) at 37ºC for 30 mins, followed by centrifugation at 4ºC to obtain the cell lysate. To immunoprecipitate CBP-GFP, 500 µg - 1 mg cell lysate was incubated with either 25 µl ChromoTek GFP-Trap® Magnetic Particles M-270 or 50 µl Protein G Dynabeads™. After recovery, beads were resuspended in 500 µl of wash buffer (NET buffer supplemented with 10mM sodium butyrate and 0.05% NP-40). 300 µl of the bead suspension was retained for a Western blot to confirm successful immunoprecipitation of the GFP-tagged protein. The remaining 200 µl was used for the in vitro HAT assay.

Beads were resuspended in HAT buffer (250 mM Tris-HCl, 25% glycerol, 0.5 mM EDTA, 250 mM KCl) supplemented with 0.1M DTT, 10mM PMSF, 1M sodium butyrate, 1mM acetyl CoA and 0.55µg/ml

recombinant polynucleosomes (H3.1; Active Motif) and the reaction mixture was incubated for 30 mins at 30ºC. After incubation, loading dye and reducing agent were added before boiling the beads. 50 µg of overall protein was loaded onto either a 3-8% Tris-Acetate gel with Tris Acetate running buffer (to visualise CBP) or a 4-12% Bis-Tris gel with MOPS running buffer (to visualise H3K27ac).

Chromatin Immunoprecipitation (ChIP):

For ChIP experiments, Dox inducible cells were plated in 10 cm dishes, 2 plates per construct. Cells were induced with concentrations of Dox that were identified using western blot as inducing CBP-GFP without altering overall CBP expression levels (Figure 4M & S4K-L). After 48h, cells were crosslinked at room temperature for 10 minutes using formaldehyde (1.1 % final). Crosslinking was quenched using 2.5 M Glycine (Melford, 125 mM final) incubating for 5 minutes at room temperature. Cells were harvested by cell scraping in PBS, and pelleted by centrifugation at 200 rcf for 5 minutes. Cell lysis was performed in ChIP lysis buffer 1 (50 mM HEPES-KOH (Sigma), pH 7.5 140 mM NaCl, 1 mM EDTA, 10 % Glycerol, 0.5 % NP40, 0.25 % Triton X - 100, Complete protease inhibitor cocktail (Roche)) by rotating at 4ºC for 5 minutes. Nuclei were pelleted by centrifugation at 1500 rcf for 5 minutes, re-suspended in ChIP Buffer 2 (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA (Sigma), Complete protease inhibitor cocktail) and incubated at room temperature on a rotator for 10 minutes. Nuclei were pelleted by centrifugation at 1500 rcf for 5 minutes and re-suspended in ChIP lysis buffer 3 (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1 % Na-deoxycholate, 0.5 % N-lauroylsarcosine (Sigma), Complete protease inhibitor cocktail) for shearing. Chromatin was sheared to 150 - 300 bp fragments, using the Bioruptor Pico (Diagenode), samples were sonicated for 12 cycles of 30 seconds on, 30 seconds off pulse sonication. Lysates were cleared by centrifugation at 20,000xg for 20 min. Concentrations of lysates was determined by BCA assay (ThermoFisher), IP's were done with 900 µg of sheared protein in a total of 300 µl with ChIP IP buffer (10 mM Tris-HCl pH 8.0, 600 mM NaCl, 1 mM EDTA, 3 % Triton X-100, Complete protease inhibitor cocktail), input reactions were made for each sample with 90 µg of protein in 50 µl of ChIP IP buffer. Immunoprecipitations were performed by incubating with 4 µg of antibody overnight with rotation at 4ºC. Immunocomplexes were recovered by adding blocked Protein G Dynabeads and incubated for 90 minutes at 4ºC with rotation. Beads were washed 5 times in ChIP wash buffer (50 mM HEPES-KOH pH 7.5, 500 mM LiCl, 1 mM EDTA, 1 % NP40, 0.7 % Na-deoxycholate, 0.1 % N-lauroylsarcosine) and 1x in ChIP final wash buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 50 mM NaCl). Sample was eluted by incubation at 65°C for 30 min in ChIP elution buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 10 mM EDTA, 1 % SDS) with rotation in a thermomixer at 900 rpm. Cross-linking was reversed by incubation at 65ºC overnight. Immunoprecipitated DNA was treated with RNase A (0.2 mg/ml final), supplemented with 4 mM $CaCl_2$ for 2 hours at 37ºC, then Proteinase K (0.2 mg/ml final) for 2 hours at 55ºC. DNA was then purified by phenol:chloroform extraction and ethanol precipitation and resuspended in 1x TE buffer.

ChIP-seq library preparation.

For ChIP-seq, samples were prepared using NEBNext Ultra™II DNA Library Prep Kit for Illumina (NEB, E7645) following manufacturers instructions. Briefly, fragmented DNA underwent 5' phosphorylation and dA-Tailing to prepare the fragment ends for adapter ligation. Sample concentration was checked using Qubit dsDNA High Sensitivity (HS) kit (Thermo Fisher), to determine adapter dilutions. For input samples the adapters were not diluted, however a 1:25 dilution of adapters in 10 mM Tris-HCl pH 7.5, was performed for immunoprecipitation samples. A cleanup of the adapter ligation was performed using AMPure XP beads (Beckman Coulter) for size selection following the 200 bp protocol. PCR amplification was performed with 3 cycles of amplification for input samples, and 13 cycles of amplification for immunoprecipitation samples. A cleanup of the PCR reaction was performed using AMPure XP beads (Beckman Coulter), following protocol guidelines. Fragment sizes were checked before sequencing using the TapeStation (Agilent), using High Sensitivity DNA ScreenTape (Agilent).

ChIP-seq data analysis:

ChIP-seq samples were sequenced using paired-end 150 sequencing on an Illumina NovaSeq 6000 platform (Novogene). Raw data quality was assessed using FastQC v0.11.9[81] and adapters were trimmed using Cutadapt v3.4 [82] implemented in Trim Galore (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Trimmed reads were aligned to human genome assembly GRCh38 using Burrows-Wheeler Alignment (BWA) v0.7.17-r1188[83]. For downstream analysis, replicates were merged and duplicate reads removed using Picard v2.27.4 (https://broadinstitute.github.io/picard/), before read quality filtering using SAMTools v1.15.1[84]. BigWig files were generated by normalizing per million reads mapped (RPM) and converting to bedgraph format with Bedtools v2.30.0[85], then to BigWig format using UCSC utilities (bedGraphToBigWig v377)[86]. ChIP-seq peaks were called on filtered data using MACS2 v2.2.7.1[60] using the BAMPE option and a broad peak cutoff of 0.001. All gene co-ordinates were obtained from Refseq. All initial processing, alignment and peak calling was completed as part of the Nextflow ChIP-seq pipeline v2.00[87] on the University of Sheffield High Performance Computing (HPC) cluster. Intersection of ChIP-seq peaks, calculation of relative distance plots[61] and Jaccard statistics[61] were done using Bedtools v2.31.0. For intersection with genome regions, intersection was made in the order Promoter
(1kb window upstream of TSS) > Exons > Introns > Intergenic. Control peaks corresponding to each sample were calculated using bedtools shuffle, over a window restricted to 40kb upstream of transcription start sites on the same chromosome. Subsequent analysis was done using a local installation of DeepTools v3.5.5 [88]. Read density coverage was calculated using multiBigwigSummary in 1kb bins over the entire genome; Correlation coefficients between samples were calculated using plotCorrelation; Heatmaps and metaplots displaying signals over a 5kb window aligned to peak centres were generated using computeMatrix and plotHeatmap; Read density enrichment over defined peak regions were calculated using plotEnrichment. All downstream analysis was done using R in Rstudio; significance p values for tag enrichment were calculated using a two-sided Mann-Whitney U-test in R[78]. All sequencing data analysed in this study have been deposited at NCBI's GEO database (https://www.ncbi.nlm.nih.gov/geo/) and assigned the identifier GSE268266.

Cleavage Under Targets and Release Using Nuclease (CUT&RUN):

For CUT&RUN experiments, doxycycline inducible CBP-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP cells were plated into 10 cm dishes and induced using the same concentrations of doxycycline as in ChIP experiments for 48 hours. 100,000 cells per reaction or input sample were harvested and DNA fragments were prepared using the CST CUT&RUN Assay kit according to the manufacturer's instructions. In brief, 100,000 cells per reaction were harvested, washed and bound to concanavalin A coated magnetic beads. Antibodies for each reaction were added (0.5 µg IgG: Rabbit (DA1E) IgG (CST, part of CUT&RUN assay kit); 1 µg H3K27ac: Histone H3K27ac (Active Motif)) and incubated overnight at 4ºC. pAG-MNase was added to each tube and incubated for 1 hour, then activated through the addiction of 3 mM $CaCl_2$ for 30 minutes at 4ºC. Stop buffer was added, supplemented with 50 pg Spike-in DNA per sample (CST, part of CUT&RUN assay kit) and DNA fragments were eluted from the beads. DNA was then purified by phenol:chloroform extraction and ethanol precipitation and resuspended in 50 µl 1x TE buffer.

CUT&RUN library preparation:

CUT&RUN DNA fragments were used to prepare libraries using the NEBNext Ultra II Library prep kit as described for ChIP-seq with the following changes. Adaptors were diluted for 1:10 for non-input samples and were cleaned up without size selection. Between 10-14 cycles of PCR were used for adaptor ligation dependent upon starting DNA concentration.

CUT&RUN data analysis:

CUT&RUN sequencing data were processed and analyzed following a spike-in normalization strategy to enable quantitative comparisons across samples. Samples were sequenced using paired-end 150 sequencing on an Illumina NovaSeq 6000 platform (Novogene). Raw data quality was assessed using FastQC v0.11.9[81] and adapters were trimmed using Cutadapt v3.4 [82] implemented in Trim Galore (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Trimmed reads were aligned to two distinct reference genomes: Homo sapiens GRCh38 (Target genome) or S.cervisiae SacCer3 R64 (spike-in control genome) using bowtie2 v2.5.4[89] in very sensitive local alignment mode (--local --very-sensitive-local --no-unal --no-mixed --no-discordant -q -I 10 -X 700 --dovetail -x), before read quality filtering using SAMTools v1.15.1[84]. We used featureCounts v2.0.6 [90] to generate a count matrix for each sample containing the total number of reads uniquely mapping to the spike-in genome. This was used to calculate a scaling factor for each sample using DEseq2 v1.40.2 [91]. To enable direct quantitative comparison of signal intensity across samples, normalized BigWig files were generated in MACS3 v3.0.0b1. Firstly we created bedgraph files using the callpeak function on MACS3, and then scaled by the previously calculated spike-in scaling factor. Normalized bedgraph files were then used to calculate the fold enrichment relative to the IgG background

control. H3K27ac enriched peaks were called using MACS3. Peak intensity values from spike-in normalized BigWig files were then used for inter-sample comparisons and differential analyses using a local installation of DeepTools v3.5.3 [88] as described previously for ChIPseq data. All sequencing data analysed in this study have been deposited at NCBI's GEO database (https://www.ncbi.nlm.nih.gov/geo/) and assigned the identifier **GSE299613 (CUT&RUN)**.

RNA-seq

RNA extraction

For RNA-seq doxycycline inducible CBP-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP cells were plated into 10 cm dishes and induced using the same concentrations of doxycycline as described previously. RNA was extracted using TRIzol and chloroform, DNase treated to remove genomic DNA, then further purified with a subsequent phenol chloroform extraction and ethanol precipitation. RNA concentrations were measured using the broad range RNA Qubit kit (ThermoFisher).

RNA-seq library preparation

RNA-seq libraries were prepared using the Quant-seq 3' mRNA-Seq Library Prep kit according to the manufacturer's instructions. In brief, 500 ng RNA was used per sample for cDNA synthesis, subsequent RNA removal and purification. qPCR was used to identify the optimal number of cycles for PCR amplification: between 16-23 depending upon sample (qPCR Add-on Kit, Lexogen) and the cDNA was then amplified accordingly, followed by purification of the libraries using magnetic beads.

RNA-seq data analysis:

RNA-seq samples were sequenced using paired-end 150 sequencing on an Illumina NovaSeq 6000 platform (Novogene). Following sequencing, Read 2 was discarded and downstream data analysis was performed using only Read 1, according to manufacturer's recommendations for libraries prepared using the Quant-seq 3' mRNA-Seq Library Prep kit. Raw data quality was assessed using FastQC v0.11.9(Babraham bioinformatics - FastQC A qu...) and adapters were trimmed using Cutadapt v3.4 (Martin 2011) implemented in Trim Galore (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Reads were aligned using STAR (v2.7.6)[92] allowing soft-clipping of read ends during alignment to generate a geneCounts matrix. This was then used as input for all downstream analysis in DESeq2 v1.40.2 in Bioconductor[91]. All sequencing data analysed in this study have been deposited at NCBI's GEO database (https://www.ncbi.nlm.nih.gov/geo/) and assigned the identifier **GSE299614 (RNA-seq)**.

Computational analysis

Identification of CBP-IDRs

To predict disordered regions in CBP, we used DISOPRED3 [93] and PONDR[37] disorder prediction software. Tested CBP-IDRs were determined using PONDR VSL2 predictor [37], where stretches of amino acids that have a prediction of disorder above 50% were classed as intrinsically disordered regions [36]. A mutant was also generated spanning the region where CBP binds to FUS [38] we termed the CBP-FUS Interaction Domain (CFID).

Analysis of CBP sequence properties

NARDINI+

Sequence-level features of the target protein were analyzed using NARDINI+ (version 1.1)[43,44] implemented through a Colab Notebook. Analysis of Full length CBP_{wt} was done using the 'NARDINI+_from_accession' iPython notebook with Uniprot accession Q92793. IDRs identified by NARDINI were subsequently mapped to our CBP-IDRs identified using PONDR[37]. In particular, 'IDR #1' from NARDINI+ was not identified by PONDR, as it was less than 50 amino acids in length; IDR #7 from NARDINI+ corresponded to CBP_{AL}. For individual IDR sequences and CBP shuffle mutant sequences, we implemented the 'NARDINI+_from_fasta' iPython notebook. The resulting quantitative z-score vectors were used to characterize the protein's molecular grammar and predict its functional roles.

localCIDER

Analysis of sequence properties was carried out using a local installation of localCIDER[50]. Linear composition plots for full length CBP_{wt} and CBP_{CFID} were calculated using a blob length window of 200bp and 50bp respectively.

PLAAC analysis

PLAAC analysis was carried out using the web-based PLAAC server[55].

Generating shuffle variants.

Shuffle sequence variants were generated in GOOSE[52] using the 'Constant residues variant' option to randomly shuffle residues without changing the overall amino acid composition of the IDRs.

AlphaFold structure prediction (version 2.3.2)

All structure predictions were performed using AlphaFold version 2.3.2[35] implemented through a Colab notebook. We used the monomer model with relaxation and specified five recycles. AlphaFold produces a per-residue confidence metric called predicted local distance difference test (pLDDT) on a scale from 0 to 100. pLDDT estimates how well the prediction would agree with an experimental structure; a pLDDT >70 is considered to correspond to a generally correct backbone prediction[94]. To assess the confidence of predictions carried out on CBP C-termnal domains (Figure S3E), we provided a duplicate view of our model coloured by pLDDT score. All visualisation was done using UCSF ChimeraX[95].

Resource availability

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Daniel Bose (d.bose@sheffield.ac.uk). All plasmids generated for this study are available from Addgene. All unique/stable reagents generated in this study are available from the lead contact without restriction. The high-throughput sequencing data data from this publication have been deposited to the GEO database (https://www.ncbi.nlm.nih.gov/geo/) and assigned the identifiers GSE268266 (ChIP-seq), GSE299613 (CUT&RUN) and GSE299614 (RNA-seq). Datasets GSE268266 GSE299613 GSE299614 have been grouped under **SuperSeries GSE299618**.

Figure legends

Figure 1: CBP forms fluid-like puncta in the nucleus.

A) Immunofluorescence of CBP in HEK293T cells; Scale bars: 10 µm.

B) Western blot for CBP and HaloTag in HEK293T cells and CBP-Halo HEK293T cells.

C) Confocal microscopy in fixed HEK293T and CBP-Halo HEK293T cells using the fluorescent Halo ligand TMR (CBP-Halo$_{TMR}$); Scale bars: 10 µm.

D) Lattice light sheet microcopy of endogenous CBP-Halo$_{TMR}$ condensates; Single plane images of endogenous CBP-Halo including merged image of Z plane, X/Z and Y/Z planes. Scale bar: 5 µm.

E) Still images from Supplementary Movie 1 showing the dynamic behaviour of CBP-Halo condensates. Boxes indicate regions enlarged in the lower panel; arrows highlight representative dynamic condensates; Scale bars: 10 µm.

F-G) FRAP of CBP-Halo condensates. F) Still images of live cell confocal microscopy showing FRAP of endogenous CBP-Halo$_{JF549}$. Scale bars: 10 µm; G) FRAP curves for endogenous CBP-Halo$_{JF549}$. Data was collected for 32 individual nuclei, and FRAP curves were calculated for 7 condensates where the bleached area did not diffuse away from the focal plane during the recovery period. Grey shading highlights imaging pre-bleach; Green line denotes bleaching. Still images from Supplementary Movie 2

H) Domain map and predicted intrinsic disorder of full length CBP$_{wt}$ using: DISOPRED3[93] (Top); PONDR[37], (bottom). Positions of CBP-IDRs (CBP$_{IDR1-7}$, CBP$_{AL}$ and CBP$_{CFID}$) with their domain boundaries predicted using PONDR, and structured CBP domains are highlighted. IDRs that were predicted but not tested are indicated in grey.

I) Live-cell confocal microscopy images showing the different behaviours of CBP$_{IDR6}$ and CBP$_{IDR7}$ in the OptoDroplet assay. Cells were imaged for 200s following exposure to blue light. Dashed box highlights the enlarged region. Scale bars: 10 µm

J) Fold change in the calculated SD of the grey values for defined individual nuclei following blue light stimulation. Diffuse signal should have a low SD, while strong changes in intensity due to condensate formation should have a high SD. Grey shading highlights the 30 second window used to calculate the initial rate of change. Bars represent mean ± s.e.m, n=3.

K) Initial rate of change over the first 30 seconds following blue light exposure. AL, Autoregulatory Loop; -ve, Opto-NLS negative control. Bars represent mean ± s.e.m; significant p values are reported (ns, not significant), n=3.

Figure S1

A-B) Analysis of PCR products from edited locus: (A) PCR using primers spanning the 3' end of CBP_{wt}; (B) Inference of CRISPR Edits (ICE)[96] analysis of Sanger sequencing data from polyclonal CBP-Halo HEK293T (Top) and HEK293T control cells (bottom).

C) RT-qPCR using primers for CBP or HaloTag in control HEK293T cells (orange) and CBP-Halo HEK293T cells (magenta). Bars represent mean ± s.e.m, n=3.

D) AlphaFold v2.3.2[35] prediction for CBP_{wt}. Domains are coloured the same as Figure 1H.

E) Analysis of sequence composition and patterning in CBP-IDRs using NARDINI+[43,44] which compares enrichment or depletion of compositional features relative to all human IDR-containing proteins. Features represented by negative z-scores are generally depleted in the IDR, while features with positive z-scores are generally enriched; features where $-1 \leq z \leq +1$ are not considered to be enriched or depleted compared to representation in all human IDRs[43,44]. Features referred to in the main text are highlighted in bold.

F) CBP-IDR constructs used for OptoDroplet assay.

G) Live-cell confocal microscopy images showing behaviour of all tested CBP-IDRs in OptoDroplet assay. Cells were imaged for 200s following exposure to blue light. Datasets for CBP_{IDR2} and CBP_{IDR5} were collected independently. Orange boxes indicate regions enlarged in bottom panel. Scale bars: 10 µm. Blue dashed box highlights data in Figure 1I.

H-I) Optodroplet analysis of CBP_{IDR2} & CBP_{IDR5}. H) Fold change in the calculated SD of the grey values. Grey shading highlights the 32 second window used to calculate the initial rate of change. +ve, Opto-FUS_{IDR} positive control; -ve, Opto-NLS negative control. Bars represent mean ± s.e.m, n=3; I) Initial rate of change over the first 32 seconds following blue light exposure. Bars represent mean ± s.e.m; significant p values are reported (ns, not significant), n=3.

Figure 2: CBP-IDRs affect the behaviour of full length CBP_{wt}

A) Schematic of $CBP_{\Delta IDR}$-GFP constructs. Deleted regions are highlighted with dashed lines. For $CBP_{\Delta AL}$-GFP, the AL is replaced with a flexible 12 residue linker (blue text).

B) Confocal microscopy in fixed transfected HEK293T cells showing the behaviour of CBP-GFP condensates. Data shows a maximum projection over the entire Z-stack. Scale bars: 10 µm.

C) Percentage of nuclei displaying diffuse or punctate signal. Images were initially selected for successful transfection based on the presence or absence of GFP signal then the data was blinded and scored on whether GFP signal was punctate or diffuse.

D) Standard deviation in grey values for nuclei transfected with CBP-GFP. p values were calculated using a Kruskal-Wallis test.

E) Number of condensates per nuclei following filtering. Dashed line represents a threshold value of 6 puncta per nuclei. Data points highlight individual nuclei from 3 biological replicates.

F) Integrated intensity density (IID) of condensates formed by $CBP_{\Delta IDR}$-GFP constructs. p values were calculated using a Kruskal-Wallis test.

G) FRAP analysis of individual CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta AL}$-GFP puncta. F) Confocal microscopy images showing photobleaching and recovery of CBP-GFP signal. Arrow indicates the bleached puncta. Scale bars: 10 µm.

H) Relative fluorescence recovery normalised to pre-bleached intensity for a 90s time period following bleaching. Grey shading highlights imaging pre-bleach; Green line denotes bleaching. Still images from Supplementary Movies 6-8.

I) Effect of 1,6 Hexanediol on CBP-GFP puncta. HEK293T cells were transfected with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta AL}$-GFP. Images were taken at fixed time points following application of 5% w/v (final concentration) 1,6 Hexanediol. Scale bars: 10 µm. Still images from Supplementary Movies 3-5.

J-K) RNAse treatment disrupts CBP condensates. Live cell images of transfected HEK293T cells with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP. Cells were treated with 100 ug/ml RNase A for 20 minutes, images were taken pre- and post-treatment. Scale bars: 5 µm

L) Lattice light sheet microscopy showing dominant phase separation phenotype on endogenous CBP condensates (JF-646) when HEK293T cells were transfected with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP. Single plane images of endogenous CBP-Halo, and transfected protein including merged image of Z plane, X/Z and Y/Z planes. Scale bar: 5 µm.

M-O) Properties of endogenous CBP puncta following transfection of CBP-GFP: CBP-Halo HEK293T cells were transfected with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP or $CBP_{\Delta IDR7}$-GFP for 24h. Endogenous CBP-Halo was labelled with Janelia Fluor® 646 Halo ligand ($CBP\text{-}Halo_{JF646}$). M) Number of condensates per nuclei, dashed line represents a threshold value of 6 puncta per nuclei; N) Integrated intensity density (IID) of condensates; O) Condensate volume in μM^3. Data points highlight individual nuclei from 3 biological replicates, p values were calculated using a Kruskal-Wallis test.

1407

1408

1409 Figure S2

1410 A) Western blot for CBP and GFP shows relative expression levels of transiently transfected CBP-GFP
1411 constructs.

1412 B) Reference images used to guide blind scoring for Figure 2C.

1413 C) Filtering strategy used for selecting transfected nuclei for analysis. Transfected nuclei were selected based
1414 on an intensity threshold [76]. GFP puncta were then filtered by: i) Area - to remove overexpression artefacts (e.g
1415 oversaturated fluorescent nuclei); ii) noise less than 3 pixels in diameter; iii) Circularity - to isolate single from
1416 overlapping puncta. Images show examples of nuclei removed by the filter, scale bars: 10 µm.

1417 D) Summary of filtering results. Highlighted populations are: Nuclei passing filter (Purple); Large area to
1418 remove oversaturated cells/nuclei with no clear puncta (dark blue); low circularity (<0.25) to remove
1419 overlapping puncta (light blue); Puncta < 3 pixels to remove speckle noise (salmon).

1420 E) Percentage of transfected nuclei for each condition remaining following application of the filter.

1421 F) Mean ± s.e.m and SD in the number of condensates per nuclei and mean IID ± s.e.m.

1422 G-I) Behaviour of $CBP_{\Delta IDR2}$-GFP and $CBP_{\Delta IDR5}$-GFP. G) Standard deviation in grey values for nuclei
1423 transfected with CBP-GFP. p values were calculated using a Kruskal-Wallis test; H) Number of condensates
1424 per nuclei following filtering. Dashed line represents a threshold value of 6 puncta per nuclei. Data points
1425 highlight individual nuclei from 3 biological replicates; I) Integrated intensity density (IID) of condensates. p
1426 values were calculated using a Kruskal-Wallis test.

1427 J) Calculated values for recovery rate over the time course following bleaching and the time to reach 50%
1428 pre-bleach intensity ($T_{1/2}$).

1429 K) Percentage of punctate nuclei following RNAse treatment.

1430 L) Number of CBP condensates per nucleus following RNAse treatment.

1431 M) CBP-IDR mutations have a dominant negative/positve effect on the behaviour of endogenous CBP
1432 condensates. CBP-Halo HEK293T cells were transfected with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP or $CBP_{\Delta IDR7}$-GFP for
1433 24h. Endogenous CBP-Halo was labelled with Janelia Fluor® 646 Halo ligand (CBP-$Halo_{JF646}$).

1434 N) Multiple channel control for live cell imaging in Figure S2J. CBP-Halo was labelled with Janelia Fluor®
1435 646 Halo ligand (CBP-$Halo_{JF646}$).

Figure 3: Sequence properties of CBP-IDRs.

A) Sequence properties of CBP_{wt}, CBP_{CFID}, CBP_{IDR6} and CBP_{IDR7}. Amino acids are grouped by property: Polar (Q,N,S,T,G,H,C); Proline (P); Aliphatic (A,L,M,I,V); Charged (E,D,R,K); Aromatic (F,Y,W) based on localCIDER defaults[50]. Numbers indicate the fraction of amino acids in each domain with the grouped property.

B) Linear amino acid composition of CBP_{CFID} calculated using localCIDER[50]. The relative positions of CBP_{IDR6} (light blue) and CBP_{IDR7} (Grey) are highlighted, along a schematic of their positions in the context of CBP_{wt} and CBP_{CFID}; The position of the 18 residue polyQ expansion is highlighted by the green bar and dashed lines. Amino acids are grouped by property: Charged (magenta); Polar (green); Aliphatic (blue); Aromatic (yellow); Proline (purple); Thin lines give the per-residue composition; Sliding window used to calculate local composition=50 residues; Stepsize between windows = 1 residue.

C) Comparison of sequence composition and patterning of CBP_{IDR6} and CBP_{IDR7} with $CBP_{IDR6\text{-}Shuffle}$ and $CBP_{IDR7\text{-}Shuffle}$ using NARDINI+[43,44]. Features represented by negative z-scores are depleted in the IDR (relative to all human IDR proteins), while features with positive z-scores are enriched; features where $-1 \leq z \leq +1$ are not considered to be enriched or depleted compared to representation in all human IDRs[43,44]. Features referred to in the main text are highlighted in bold.

D-F) Effect of Shuffled IDRs on CBP condensates. D) Fixed cell confocal microscopy comparing CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP to their shuffled counterpart. Scale bars: 10 µm; E) Number of condensates per nuclei following filtering. Dashed line represents a threshold value of 6 puncta per nuclei. Data points highlight individual nuclei from 3 biological replicates; F) Integrated intensity density (IID) of condensates. p values were calculated using a Kruskal-Wallis test.

G) PLAAC analysis[53,55] of CBP's C-terminal domains (residues 1856-2441), showing CBP_{wt}, $CBP_{IDR6\text{-}shuffle}$ and $CBP_{IDR7\text{-}shuffle}$. The relative positions of CBP_{IDR6} and CBP_{IDR7} are highlighted.

Figure S3

A-B) Randomly shuffled sequences for: A) $CBP_{IDR6\text{-}shuffle}$ and; B) $CBP_{IDR7\text{-}shuffle}$ calculated using GOOSE[52].

C) Percentage of transfected nuclei for each condition remaining after filtering.

D) Percentage of punctate nuclei CBP_{wt}, $CBP_{IDR6\text{-}shuffle}$ and $CBP_{IDR7\text{-}shuffle}$.

E) AlphaFold v2.3.2[35] prediction of CBP's C-terminal domains. Left panel: The positions of the predicted interacting α-helices CBP_{IDR6H2} (cyan), CBP_{IDR7H1} (dark blue), CBP_{IDR7H2} (light blue), and the CBP_{NCBD} are highlighted; Right Panel) pLDDT confidence scores.

F) Domain structure of the CBP_{CFID} highlighting the positions of CBP_{IDR6} (cyan) and CBP_{IDR7} (light blue). CBP_{NCBD} and predicted helices are highlighted. Dashed lines show the interactions between α-helices predicted by AlphaFold v2.3.2[35].

G) Fixed cell confocal microscopy comparing the behaviour of CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP with $CBP_{\Delta IDR6H2}$-GFP and $CBP_{\Delta IDR7H1}$-GFP. Scale bars: 10 µm.

H-I) Comparison of condensates formed by $CBP_{\Delta IDR6H2}$-GFP and $CBP_{\Delta IDR7H1}$-GFP with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP. H) Number of puncta per nuclei; I) Integrated intensity density. p values were calculated using a Kruskal-Wallis test.

J) Amino acid sequence conservation between CBP and p300 from local pairwise sequence alignment using EMBOSS Matcher[97]. N-term, regions N-terminal to the HAT domain residues CBP 1-1195; p300 1-1159. (Full alignments are available in Supplementary material).

K) Comparison of sequence composition and patterning of CBP_{IDR6} and CBP_{IDR7} with $p300_{IDR6}$ and $p300_{IDR7}$ using NARDINI+[43,44]. Features represented by negative z-scores are depleted in the IDR (relative to all human IDR proteins), while features with positive z-scores are enriched; features where $-1 \leq z \leq +1$ are not considered to be enriched or depleted compared to representation in all human IDRs[43,44]. Features referred to in the main text are highlighted in bold.

L) PLAAC analysis[53,55] of CBP (purple) and p300 (Magenta) C-terminal regions. The relative positions of CBP_{IDR6} and CBP_{IDR7} are highlighted. Positions in p300 are based on local pairwise sequence alignment using EMBOSS Matcher[97] (Full alignments are available in Supplementary material).

Figure 4: CBP-IDRs affect response to acetylation

A) Fixed cell confocal microscopy of CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP after 48h treatment with 0.5uM A-485, DMSO or no treatment. Scale bars: 5 µm.

B-D) Effect of A-485 or Trichostatin A (TSA) on CBP condensate behaviour: B) Percentage of nuclei displaying diffuse or punctate signal; C) Number of condensates per nuclei following filtering. Dashed line represents a threshold value of 6 puncta per nuclei. Data points highlight individual nuclei from 3 biological replicates; D) Integrated intensity density (IID) of condensates formed by $CBP_{\Delta IDR}$-GFP constructs. Cells were treated with A-485 for 36 hours[57] and with TSA for 90 minutes[58].

E-G) A-485 concentrations affect CBP condensate behaviour: E) Fixed cell confocal microscopy of CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP after 60 minutes treatment with stated concentration of A-485 (0um = DMSO). Scale bars: 5 µm; F) Number of condensates per nuclei; G) Integrated intensity density (IID) of condensates. Scale: Pseudo-log base 10.

H-L) CBP-IDRs affect CBP's response to acetylation on the AL: H) Fixed cell confocal microscopy of CBP_{AL} double mutant constructs. Scale bars: 10 µm; I) Percentage of nuclei displaying diffuse or punctate signal; J) Number of condensates per nucleus; K) Integrated intensity density (CBP_{wt}-GFP background); L) Integrated intensity density ($CBP_{\Delta IDR6}$-GFP background).

M) Confocal microscopy showing inducible CBP-GFP expression. Stable inducible CBP-GFP HEK293T cells were treated with Doxycycline (Dox) for 48h. Dox concentrations: CBP_{wt}-GFP, 25ngml-1; $CBP_{\Delta IDR6}$-GFP, 5ngml-1; and CBP_{IDR7}-GFP, 25ngml-1; Scale bars: 10 µm.

N) Western blot of CBP, K1535ac in the AL of CBP/p300 and H3K27ac and following the addition of Doxycycline (Figure 4M & S4K) to induce expression of CBP-GFP. CBP was resolved using 3-8% Tris-Acetate PAGE (loading control: GAPDH, imaged using chemiluminescence) and H3K27ac was resolved using 12% Bis-Tris PAGE (loading control: Histone H4).

O) Immunoprecipitation HAT assay using recombinant di-nucleosome substrate. CBP-GFP expression was induced at equivalent levels using Doxycycline as in (Figure 4M & S4K) and CBP-GFP was immunoprecipitated using GFP-Trap® Magnetic beads. The reaction was resolved by 12% Bis-Tris PAGE and probed for H3K27ac and H4 by western blot.

All p-values for IID were calculated using Kruskal-Wallis tests.

Figure S4

A) Western blot of CBP, GFP, K1535ac in the AL of CBP/p300 and H3K27ac and following treatment with 500nM A-485 for 36 hours[57] and 500nM Trichostatin A (TSA)[58] for 90 minutes. CBP was resolved using 3-8% Tris-Acetate PAGE (loading control: $GAPDH^{A}$) and H3K27ac was resolved using 12% Bis-Tris PAGE (loading control: $GAPDH^{B}$).

B) Fixed cell confocal microscopy of CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP after 48h treatment with YYnM Trichostatin A (TSA), DMSO or no treatment. Scale bars: 5 µm.

C-E) Effect of A-485 concentration on CBP condensate behaviour: C) Fixed cell confocal microscopy of CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP after 120 minutes treatment with stated concentration of A-485 (0 µm = DMSO). Scale bars: 5 µm; D) Number of condensates per nuclei; E) Integrated intensity density of condensates. Scale: Pseudo-log base 10.

F) CBP_{AL} constructs showing sequences of CBP_{wt}, $CBP_{\Delta AL}$-GFP and CBP_{KTG}-GFP. In $CBP_{\Delta AL}$-GFP the deleted region was replaced by a 12-residue flexible linker (blue); Lysine residues subject to acetylation are highlighted (red, bold); positions of K>G mutations are highlighted (orange, bold).

G) Fixed cell confocal microscopy of CBP_{AL} double mutant constructs. Scale bars: 10 µm

H-I) Number of condensates per nucleus for $CBP_{\Delta AL}$-GFP and CBP_{KTG}-GFP on a background of: H) $CBP_{\Delta IDR7}$-GFP and I) $CBP_{\Delta CFID}$-GFP.

J)

K) Inducible CBP-GFP expression: CBP_{wt}-GFP HEK293T cells were treated with 0-50ng/ml Doxycycline for 48h. Confocal microscopy images. Scale bars: 10µm;

L) Inducible CBP-GFP expression: western blot for CBP and GFP. The concentration of Doxycycline that induced even CBP-GFP expression is highlighted (red)

M-O) Behaviour of CBP-GFP condensates following Dox induction. L) Percentage of nuclei displaying diffuse or punctate signal; N) Number of condensates per nuclei; O) Integrated intensity density of condensates formed by $CBP_{\Delta IDR}$-GFP constructs. *All p-values for IID were calculated using Kruskal-Wallis tests.*

P-Q) Quantification of western blots in Figure 4N: O) Foldchange in intensity of H3K27ac following normalisation to H4; P) Foldchange in intensity of CBP (K1535ac) following normalisation to CBP. (n=1).

R) Quantification of western blots from immunoprecipitation HAT assay in Figure 4O. Intensity of H3K27ac following normalisation to H4 and CBP loading controls.

Figure 5: CBP-IDRs regulate nuclear chromatin association and H3K27ac.

A) Overlap between CBP binding sites identified from CBP and GFP ChIP-seq. $CBP_{wt\text{-}Tet3G}$ (purple), CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) and CBP_{IDR7}-GFP (dark blue). Magenta bar highlights consensus peaks found in every CBP-GFP dataset.

B) Peak widths of CBP binding sites. Violin plots show CBP peak widths in $HEK293T_{Tet3G}$ cells (purple), and GFP ChIP-seq reads in CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) and CBP_{IDR7}-GFP (dark blue). p values from Mann-Whitney U-test.

C) Distribution of CBP binding sites by genome region.

D) Genome browser views of CBP and GFP ChIP-seq (top) and H3K27ac CUT&RUN (bottom) at the *TSC22D3* gene. $CBP_{wt\text{-}Tet3G}$ (Purple), CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) CBP_{IDR7}-GFP (dark blue) and H3K27ac (red). Solid bars show binding sites identified using MACS2.

E) Peak widths of enriched H3K27ac domains. Violin plots show H3K27ac peak widths following induction of CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) and CBP_{IDR7}-GFP (dark blue). p values from Mann-Whitney U-test.

F-G) Read density enrichment at H3K27ac domains. E) H3K27ac; F) CBP-GFP. H3K27ac peaks were randomly downsampled to match the size of the smallest dataset (CBP_{wt}-GFP, 9213); p values from Mann-Whitney U test.

H-I) Read density enrichment at 4,657 CBP peaks common in all datasets. E) H3K27ac; F) CBP-GFP; p values from Mann-Whitney U test.

J) Distribution of consensus and non-consensus CBP binding sites by genome region.

K-L) Heatmaps of sites unique to each CBP-IDR mutation. K) CBP-GFP ChIP-seq signal; L) H3K27ac CUT&RUN signal. Heatmaps show ±2.5 Kb window from centre of peaks; CBP_{wt}-GFP (3,049 sites, orange), $CBP_{\Delta IDR6}$-GFP (13,776 sites, light blue) CBP_{IDR7}-GFP (12,605 sites, dark blue).

M-N) Enrichment of M) CBP-GFP and N) H3K27ac read densities at 2,348,854 GRCh38 cCREs from ENCODE v4 [62] . p-values from Mann-Whitney U test.

Figure S5: CBP-IDRs regulate nuclear chromatin association and H3K27ac.

A-B) Correlation between ChIP-seq datasets: A) Pearson's correlation coefficient. Numbers denote 95% Confidence Interval (Upper and lower bounds); Degrees of freedom = 1,416,090; $p < 2.2e^{-16}$ for all values; B) R^2.

C-D) Read density enrichment at endogenous 14,913 endogenous CBP_{wt} peaks identified in CBP ChIP-seq from $CBP_{wt\text{-}Tet3G}$ cells. C) Heatmaps showing a ±2.5-kb window from the centre of each called CBP peak. D) Boxplots show CBP ChIP-seq reads in $HEK293T_{Tet3G}$ cells (purple); GFP ChIP-seq reads in CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) and CBP_{IDR7}-GFP (dark blue). p values from Mann-Whitney U-test.

E) Relative distance plots for CBP-GFP ChIP-seq peaks (GFP ChIP). When intervals are closer than expected by chance, distributions shift towards lower relative distance values. Shuffled control regions (green) were restricted to TSS -40 kb.

F) Venn diagrams showing overlap between CBP binding sites in maternal $HEK293T_{Tet3G}$ lines (purple) and CBP-GFP cell lines; CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) and CBP_{IDR7}-GFP (dark blue). p values from permutation test with random regions restricted to TSS -40 kb; Jaccard statistic (J) measures overlap between two sets of intervals; perfect correlation = 1.

G-H) Read density enrichment at 7,661 CBP_{wt}-GFP peaks. E) H3K27ac; F) CBP-GFP. $CBP_{\Delta IDR6}$-GFP (light blue) and CBP_{IDR7}-GFP sites were randomly downsampled to 7,661 peaks; p values from Mann-Whitney U test.

1592 H-I) Heatmaps of 4,657 common CBP peaks. Heatmaps show ±2.5 Kb window from the centre of peaks. H)
1593 CBP-GFP ChIP-seq; I) H3K27ac CUT&RUN.

1594 J-K) Read density enrichment at CBP binding sites unique to each CBP-IDR mutation. J) CBP-GFP; K)
1595 H3K27ac; p values from Mann-Whitney U test.

1596

1597 Figure 6: CBP-IDRs regulate gene expression profiles

1598 A-B) Differentially expressed genes between A) CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP; and B) CBP_{wt}-GFP and
1599 $CBP_{\Delta IDR7}$-GFP. Genes showing significant changes in expression using more stringent significance parameters
1600 $p < 0.001$ and $log_2FC > 3$ are highlighted (red points, gene names & Supplementary Table 4). HAND1 (bold)
1601 is upregulated in both $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP datasets.

1602 C) Genome browser views of GFP ChIP-seq (top) and H3K27ac CUT&RUN (bottom) at the *HAND1* gene.
1603 CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) CBP_{IDR7}-GFP (dark blue) and H3K27ac (red). Solid bars show
1604 identified binding sites.

1605

1606

1607 Figure S6

1608 A) Pearson correlation coefficient between RNA-seq datasets from inducible CBP-GFP HEK293T cell lines
1609 following induction of CBP-GFP expression using doxycycline (Figure 4M & S4K). Data shows two
1610 biological replicates (Rep1/Rep2) for each CBP-IDR deletion.

1611 B) Principle Components Analysis (PCA) of all RNAseq datasets.

1612 C-D) Gene expression differences between: C) CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP; and D) CBP_{wt}-GFP and
1613 $CBP_{\Delta IDR7}$-GFP. Genes showing significant changes in expression ($p<0.05$) are highlighted (red points).

1614

1615 Figure 7: Summary of CBP-IDR functions

1616 A) Summary of how CBP-IDRs regulate condensates, chromatin localisation and histone acetylation.

1617 B) Summary of how CBP-IDRs work with lysine acetylation to control condensate behaviour.

1618 Figure S7

1619 B) Summary of the contribution of CBP-IDRs to CBP condensate behaviours relative to CBP_{wt}-GFP. IDRs are
1620 clustered according to whether they are: N-terminal to the HAT domain; in the Autoregulatory Loop (AL); or
1621 C-terminal to the HAT domain. All behaviours refer to behaviour of CBP following deletion of each IDR,
1622 except the OptoDroplet data which is the behaviour of the IDR in isolation. Endogenous = behaviour of
1623 endogenous CBP condensates following transfection with CBP-GFP construct.

1624

References

1. Holehouse, A.S., and Kragelund, B.B. (2023). The molecular basis for cellular function of intrinsically disordered protein regions. Nat. Rev. Mol. Cell Biol., 1–25. https://doi.org/10.1038/s41580-023-00673-0.

2. Lobel, J.H., and Ingolia, N.T. (2025). Deciphering disordered regions controlling mRNA decay in high-throughput. Nature, 1–9. https://doi.org/10.1038/s41586-025-08919-x.

3. Moses, D., Guadalupe, K., Yu, F., Flores, E., Perez, A.R., McAnelly, R., Shamoon, N.M., Kaur, G., Cuevas-Zepeda, E., Merg, A.D., et al. (2024). Structural biases in disordered proteins are prevalent in the cell. Nat. Struct. Mol. Biol. https://doi.org/10.1038/s41594-023-01148-8.

4. Cuevas-Velazquez, C.L., Vellosillo, T., Guadalupe, K., Schmidt, H.B., Yu, F., Moses, D., Brophy, J.A.N., Cosio-Acosta, D., Das, A., Wang, L., et al. (2021). Intrinsically disordered protein biosensor tracks the physical-chemical effects of osmotic stress on cells. Nat. Commun. *12*, 5438. https://doi.org/10.1038/s41467-021-25736-8.

5. Tie, F., Banerjee, R., Stratton, C.A., Prasad-Sinha, J., Stepanik, V., Zlobin, A., Diaz, M.O., Scacheri, P.C., and Harte, P.J. (2009). CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. Development *136*, 3131–3141.

6. Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. U. S. A. *107*, 21931–21936. https://doi.org/10.1073/pnas.1016071107.

7. Tropberger, P., Pott, S., Keller, C., Kamieniarz-Gdula, K., Caron, M., Richter, F., Li, G., Mittler, G., Liu, E.T., Bühler, M., et al. (2013). Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer. Cell *152*, 859–872. https://doi.org/10.1016/j.cell.2013.01.032.

8. Narita, T., Higashijima, Y., Kilic, S., Liebner, T., Walter, J., and Choudhary, C. (2023). Acetylation of histone H2B marks active enhancers and predicts CBP/p300 target genes. Nat. Genet. *55*, 679–692. https://doi.org/10.1038/s41588-023-01348-4.

9. Weinert, B.T., Narita, T., Satpathy, S., Srinivasan, B., Hansen, B.K., Schölz, C., Hamilton, W.B., Zucconi, B.E., Wang, W.W., Liu, W.R., et al. (2018). Time-resolved analysis reveals rapid dynamics and broad scope of the CBP/p300 acetylome. Cell *174*, 231–244.e12. https://doi.org/10.1016/j.cell.2018.04.033.

10. Boija, A., Mahat, D.B., Zare, A., Holmqvist, P.-H., Philip, P., Meyers, D.J., Cole, P.A., Lis, J.T., Stenberg, P., and Mannervik, M. (2017). CBP Regulates Recruitment and Release of Promoter-Proximal RNA Polymerase II. Mol. Cell *68*, 491–503.e5. https://doi.org/10.1016/j.molcel.2017.09.031.

11. Ma, L., Gao, Z., Wu, J., Zhong, B., Xie, Y., Huang, W., and Lin, Y. (2021). Co-condensation between transcription factor and coactivator p300 modulates transcriptional bursting kinetics. Mol. Cell *81*, 1682–1697.e7. https://doi.org/10.1016/j.molcel.2021.01.031.

12. Hunt, G., Boija, A., and Mannervik, M. (2022). p300/CBP sustains Polycomb silencing by non-enzymatic functions. Mol. Cell *82*, 3580–3597.e9. https://doi.org/10.1016/j.molcel.2022.09.005.

13. Trojanowski, J., Frank, L., Rademacher, A., Mücke, N., Grigaitis, P., and Rippe, K. (2022). Transcription activation is enhanced by multivalent interactions independent of phase separation. Mol. Cell *82*, 1878–1893.e10. https://doi.org/10.1016/j.molcel.2022.04.017.

14. Gomez-Lamarca, M.J., Falo-Sanjuan, J., Stojnic, R., Abdul Rehman, S., Muresan, L., Jones, M.L., Pillidge, Z., Cerda-Moya, G., Yuan, Z., Baloul, S., et al. (2018). Activation of the notch signaling pathway in vivo elicits changes in CSL nuclear dynamics. Dev. Cell *44*, 611–623.e7. https://doi.org/10.1016/j.devcel.2018.01.020.

15. Demarest, S.J., Martinez-Yamout, M., Chung, J., Chen, H., Xu, W., Dyson, H.J., Evans, R.M., and Wright, P.E. (2002). Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. Nature *415*, 549–553.

16. Kjaergaard, M., Andersen, L., Nielsen, L.D., and Teilum, K. (2013). A folded excited state of ligand-free nuclear coactivator binding domain (NCBD) underlies plasticity in ligand recognition. Biochemistry *52*, 1686–1693. https://doi.org/10.1021/bi4001062.

17. Lee, C.W., Martinez-Yamout, M.A., Dyson, H.J., and Wright, P.E. (2010). Structure of the p53 transactivation domain in complex with the nuclear receptor coactivator binding domain of CREB binding protein. Biochemistry *49*, 9964–9971. https://doi.org/10.1021/bi1012996.

18. Lin, C.H., Hare, B.J., Wagner, G., Harrison, S.C., Maniatis, T., and Fraenkel, E. (2001). A small domain of CBP/p300 binds diverse proteins: solution structure and functional studies. Mol. Cell *8*, 581–590. https://doi.org/10.1016/s1097-2765(01)00333-1.

19. Dyson, H.J., and Wright, P.E. (2016). Role of Intrinsic Protein Disorder in the Function and Interactions of the Transcriptional Coactivators CREB-binding Protein (CBP) and p300. J. Biol. Chem. *291*, 6714–6722. https://doi.org/10.1074/jbc.R115.692020.

20. Papachristou, E.K., Kishore, K., Holding, A.N., Harvey, K., Roumeliotis, T.I., Chilamakuri, C.S.R., Omarjee, S., Chia, K.M., Swarbrick, A., Lim, E., et al. (2018). A quantitative mass spectrometry-based approach to monitor the dynamics of endogenous chromatin-associated protein complexes. Nat. Commun. *9*, 2311. https://doi.org/10.1038/s41467-018-04619-5.

21. Thompson, P.R., Wang, D., Wang, L., Fulco, M., Pediconi, N., Zhang, D., An, W., Ge, Q., Roeder, R.G., Wong, J., et al. (2004). Regulation of the p300 HAT domain via a novel activation loop. Nat. Struct. Mol. Biol. *11*, 308–315. https://doi.org/10.1038/nsmb740.

22. Bose, D.A., Donahue, G., Reinberg, D., Shiekhattar, R., Bonasio, R., and Berger, S.L. (2017). RNA binding to CBP stimulates histone acetylation and transcription. Cell *168*, 135–149.e22. https://doi.org/10.1016/j.cell.2016.12.020.

23. Delvecchio, M., Gaucher, J., Aguilar-Gurrieri, C., Ortega, E., and Panne, D. (2013). Structure of the p300 catalytic core and implications for chromatin targeting and HAT regulation. Nat. Struct. Mol. Biol. https://doi.org/10.1038/nsmb.2642.

24. Ortega, E., Rengachari, S., Ibrahim, Z., Hoghoughi, N., Gaucher, J., Holehouse, A.S., Khochbin, S., and Panne, D. (2018). Transcription factor dimerization activates the p300 acetyltransferase. Nature *562*, 538–544. https://doi.org/10.1038/s41586-018-0621-1.

25. Hou, T.Y., and Kraus, W.L. (2022). Analysis of estrogen-regulated enhancer RNAs identifies a functional motif required for enhancer assembly and gene expression. Cell Rep. *39*, 110944. https://doi.org/10.1016/j.celrep.2022.110944.

26. Carullo, N.V.N., and Day, J.J. (2019). Genomic Enhancers in Brain Health and Disease. Genes *10*. https://doi.org/10.3390/genes10010043.

27. Thompson, P.R., Kurooka, H., Nakatani, Y., and Cole, P.A. (2001). Transcriptional coactivator protein p300. Kinetic characterization of its histone acetyltransferase activity. J. Biol. Chem. *276*, 33721–33729. https://doi.org/10.1074/jbc.M104736200.

28. Zhang, Y., Brown, K., Yu, Y., Ibrahim, Z., Zandian, M., Xuan, H., Ingersoll, S., Lee, T., Ebmeier, C.C., Liu, J., et al. (2021). Nuclear condensates of p300 formed though the structured catalytic core can act as a storage pool of p300 with reduced HAT activity. Nat. Commun. *12*, 4618. https://doi.org/10.1038/s41467-021-24950-8.

29. Perissi, V., Dasen, J.S., Kurokawa, R., Wang, Z., Korzus, E., Rose, D.W., Glass, C.K., and Rosenfeld, M.G. (1999). Factor-specific modulation of CREB-binding protein acetyltransferase activity. Proc. Natl.

Acad. Sci. U. S. A. *96*, 3652–3657.

30. Ferrie, J.J., Karr, J.P., Tjian, R., and Darzacq, X. (2022). “Structure”-function relationships in eukaryotic transcription factors: The role of intrinsically disordered regions in gene regulation. Mol. Cell. https://doi.org/10.1016/j.molcel.2022.09.021.

31. LaMorte, V.J., Dyck, J.A., Ochs, R.L., and Evans, R.M. (1998). Localization of nascent RNA and CREB binding protein with the PML-containing nuclear body. Proc. Natl. Acad. Sci. U. S. A. *95*, 4991–4996. https://doi.org/10.1073/pnas.95.9.4991.

32. Girdwood, D., Bumpass, D., Vaughan, O.A., Thain, A., Anderson, L.A., Snowden, A.W., Garcia-Wilson, E., Perkins, N.D., and Hay, R.T. (2003). P300 transcriptional repression is mediated by SUMO modification. Mol. Cell *11*, 1043–1054.

33. Boisvert, F.M., Kruhlak, M.J., Box, A.K., Hendzel, M.J., and Bazett-Jones, D.P. (2001). The transcription coactivator CBP is a dynamic component of the promyelocytic leukemia nuclear body. J. Cell Biol. *152*, 1099–1106. https://doi.org/10.1083/jcb.152.5.1099.

34. Nucifora, F.C., Sasaki, M., Peters, M.F., Huang, H., Cooper, J.K., Yamada, M., Takahashi, H., Tsuji, S., Troncoso, J., Dawson, V.L., et al. (2001). Interference by huntingtin and atrophin-1 with cbp-mediated transcription leading to cellular toxicity. Science *291*, 2423–2428.

35. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.

36. Basu, S., Mackowiak, S.D., Niskanen, H., Knezevic, D., Asimi, V., Grosswendt, S., Geertsema, H., Ali, S., Jerković, I., Ewers, H., et al. (2020). Unblending of Transcriptional Condensates in Human Repeat Expansion Disease. Cell. https://doi.org/10.1016/j.cell.2020.04.018.

37. Xue, B., Dunbrack, R.L., Williams, R.W., Dunker, A.K., and Uversky, V.N. (2010). PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. Biochim. Biophys. Acta *1804*, 996–1010. https://doi.org/10.1016/j.bbapap.2010.01.011.

38. Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., Tempst, P., Rosenfeld, M.G., Glass, C.K., and Kurokawa, R. (2008). Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. Nature *454*, 126–130.

39. Shin, Y., Berry, J., Pannucci, N., Haataja, M.P., Toettcher, J.E., and Brangwynne, C.P. (2017). Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets. Cell *168*, 159–171.e14. https://doi.org/10.1016/j.cell.2016.11.054.

40. Luo, F., Gui, X., Zhou, H., Gu, J., Li, Y., Liu, X., Zhao, M., Li, D., Li, X., and Liu, C. (2018). Atomic structures of FUS LC domain segments reveal bases for reversible amyloid fibril formation. Nat. Struct. Mol. Biol. *25*, 341–346. https://doi.org/10.1038/s41594-018-0050-8.

41. Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T.M., et al. (2015). A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. Cell *162*, 1066–1077. https://doi.org/10.1016/j.cell.2015.07.047.

42. Park, S., Stanfield, R.L., Martinez-Yamout, M.A., Dyson, H.J., Wilson, I.A., and Wright, P.E. (2017). Role of the CBP catalytic core in intramolecular SUMOylation and control of histone H3 acetylation. Proceedings of the National Academy of Sciences *114*, E5335–E5342.

43. King, M.R., Ruff, K.M., Lin, A.Z., Pant, A., Farag, M., Lalmansingh, J.M., Wu, T., Fossat, M.J., Ouyang, W., Lew, M.D., et al. (2024). Macromolecular condensation organizes nucleolar sub-phases to set up a pH gradient. Cell *187*, 1889–1906.e24. https://doi.org/10.1016/j.cell.2024.02.029.

44. Ruff, K.M., King, M.R., Ying, A.W., Liu, V., Pant, A., Lieberman, W.E., Shinn, M.K., Su, X., Kadoch,

C., and Pappu, R.V. (2025). Molecular grammars of predicted intrinsically disordered regions that span the human proteome. Cell *0*. https://doi.org/10.1016/j.cell.2025.10.019.

45. Rodríguez, C.S., and Laurents, D.V. (2024). Architectonic principles of polyproline II helix bundle protein domains. Arch. Biochem. Biophys. *756*, 109981. https://doi.org/10.1016/j.abb.2024.109981.

46. Piai, A., Calçada, E.O., Tarenzi, T., Grande, A.D., Varadi, M., Tompa, P., Felli, I.C., and Pierattelli, R. (2016). Just a Flexible Linker? The Structural and Dynamic Properties of CBP-ID4 Revealed by NMR Spectroscopy. Biophys. J. *110*, 372–381. https://doi.org/10.1016/j.bpj.2015.11.3516.

47. Chong, S., Graham, T.G.W., Dugast-Darzacq, C., Dailey, G.M., Darzacq, X., and Tjian, R. (2022). Tuning levels of low-complexity domain interactions to modulate endogenous oncogenic transcription. Mol. Cell *82*, 2084–2097.e5. https://doi.org/10.1016/j.molcel.2022.04.007.

48. Chong, S., Dugast-Darzacq, C., Liu, Z., Dong, P., Dailey, G.M., Cattoglio, C., Heckert, A., Banala, S., Lavis, L., Darzacq, X., et al. (2018). Imaging dynamic and selective low-complexity domain interactions that control gene transcription. Science *361*, eaar2555. https://doi.org/10.1126/science.aar2555.

49. Riback, J.A., Zhu, L., Ferrolino, M.C., Tolbert, M., Mitrea, D.M., Sanders, D.W., Wei, M.-T., Kriwacki, R.W., and Brangwynne, C.P. (2020). Composition-dependent thermodynamics of intracellular phase separation. Nature. https://doi.org/10.1038/s41586-020-2256-2.

50. Holehouse, A.S., Das, R.K., Ahad, J.N., Richardson, M.O.G., and Pappu, R.V. (2017). CIDER: Resources to analyze sequence-ensemble relationships of intrinsically disordered proteins. Biophys. J. *112*, 16–21. https://doi.org/10.1016/j.bpj.2016.11.3200.

51. Martin, E.W., Holehouse, A.S., Grace, C.R., Hughes, A., Pappu, R.V., and Mittag, T. (2016). Sequence Determinants of the Conformational Properties of an Intrinsically Disordered Protein Prior to and upon Multisite Phosphorylation. J. Am. Chem. Soc. *138*, 15323–15335. https://doi.org/10.1021/jacs.6b10272.

52. Emenecker, R.J., Guadalupe, K., Shamoon, N.M., Sukenik, S., and Holehouse, A.S. (2023). Sequence-ensemble-function relationships for disordered proteins in live cells. bioRxivorg, 2023.10.29.564547. https://doi.org/10.1101/2023.10.29.564547.

53. Alberti, S., Halfmann, R., King, O., Kapila, A., and Lindquist, S. (2009). A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. Cell *137*, 146–158. https://doi.org/10.1016/j.cell.2009.02.044.

54. Wang, J., Choi, J.-M., Holehouse, A.S., Lee, H.O., Zhang, X., Jahnel, M., Maharana, S., Lemaitre, R., Pozniakovsky, A., Drechsel, D., et al. (2018). A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. Cell *174*, 688–699.e16. https://doi.org/10.1016/j.cell.2018.06.006.

55. Lancaster, A.K., Nutter-Upham, A., Lindquist, S., and King, O.D. (2014). PLAAC: a web and command-line application to identify proteins with prion-like amino acid composition. Bioinformatics *30*, 2501–2502. https://doi.org/10.1093/bioinformatics/btu310.

56. Ibrahim, Z., Wang, T., Destaing, O., Salvi, N., Hoghoughi, N., Chabert, C., Rusu, A., Gao, J., Feletto, L., Reynoird, N., et al. (2022). Structural insights into p300 regulation and acetylation-dependent genome organisation. Nat. Commun. *13*, 7759. https://doi.org/10.1038/s41467-022-35375-2.

57. Lasko, L.M., Jakob, C.G., Edalji, R.P., Qiu, W., Montgomery, D., Digiammarino, E.L., Hansen, T.M., Risi, R.M., Frey, R., Manaves, V., et al. (2017). Discovery of a selective catalytic p300/CBP inhibitor that targets lineage-specific tumours. Nature *550*, 128–132. https://doi.org/10.1038/nature24028.

58. Yi, F., Cai, C., Ruan, B., Hao, M., Yeo, S.K., Haas, M., Yang, F., Zhang, X., and Guan, J.-L. (2023). Regulation of RB1CC1/FIP200 stability and autophagy function by CREBBP-mediated acetylation in an intrinsically disordered region. Autophagy *19*, 1662–1677. https://doi.org/10.1080/15548627.2022.2148432.

59. Henninger, J.E., Oksuz, O., Shrinivas, K., Sagi, I., LeRoy, G., Zheng, M.M., Andrews, J.O., Zamudio, A.V., Lazaris, C., Hannett, N.M., et al. (2020). RNA-mediated feedback control of transcriptional condensates. Cell. https://doi.org/10.1016/j.cell.2020.11.030.

60. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137. https://doi.org/10.1186/gb-2008-9-9-r137.

61. Favorov, A., Mularoni, L., Cope, L.M., Medvedeva, Y., Mironov, A.A., Makeev, V.J., and Wheelan, S.J. (2012). Exploring massive, genome scale datasets with the GenometriCorr package. PLoS Comput. Biol. *8*, e1002529. https://doi.org/10.1371/journal.pcbi.1002529.

62. ENCODE Project Consortium, Moore, J.E., Purcaro, M.J., Pratt, H.E., Epstein, C.B., Shoresh, N., Adrian, J., Kawli, T., Davis, C.A., Dobin, A., et al. (2020). Expanded encyclopaedias of DNA elements in the human and mouse genomes. Nature *583*, 699–710. https://doi.org/10.1038/s41586-020-2493-4.

63. He, J., Huo, X., Pei, G., Jia, Z., Yan, Y., Yu, J., Qu, H., Xie, Y., Yuan, J., Zheng, Y., et al. (2024). Dual-role transcription factors stabilize intermediate expression levels. Cell. https://doi.org/10.1016/j.cell.2024.03.023.

64. Boeynaems, S., Dorone, Y., Zhuang, Y., Shabardina, V., Huang, G., Marian, A., Kim, G., Sanyal, A., Şen, N.-E., Griffith, D., et al. (2023). Poly(A)-binding protein is an ataxin-2 chaperone that regulates biomolecular condensates. Mol. Cell. https://doi.org/10.1016/j.molcel.2023.05.025.

65. Ferrie, J.J., Karr, J.P., Graham, T.G.W., Dailey, G.M., Zhang, G., Tjian, R., and Darzacq, X. (2024). p300 is an obligate integrator of combinatorial transcription factor inputs. Mol. Cell *84*, 234–243.e4. https://doi.org/10.1016/j.molcel.2023.12.004.

66. Xu, L., Xuan, H., He, W., Zhang, L., Huang, M., Li, K., Wen, H., Xu, H., and Shi, X. (2023). TAZ2 truncation confers overactivation of p300 and cellular vulnerability to HDAC inhibition. Nat. Commun. *14*. https://doi.org/10.1038/s41467-023-41245-2.

67. Petrij, F., Giles, R.H., Dauwerse, H.G., Saris, J.J., Hennekam, R.C., Masuno, M., Tommerup, N., van Ommen, G.J., Goodman, R.H., and Peters, D.J. (1995). Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. Nature *376*, 348–351. https://doi.org/10.1038/376348a0.

68. Wang, J., Weaver, I.C.G., Gauthier-Fisher, A., Wang, H., He, L., Yeomans, J., Wondisford, F., Kaplan, D.R., and Miller, F.D. (2010). CBP histone acetyltransferase activity regulates embryonic neural differentiation in the normal and Rubinstein-Taybi syndrome brain. Dev. Cell *18*, 114–125. https://doi.org/10.1016/j.devcel.2009.10.023.

69. Haghshenas, S., Bout, H.J., Schijns, J.M., Levy, M.A., Kerkhof, J., Bhai, P., McConkey, H., Jenkins, Z.A., Williams, E.M., Halliday, B.J., et al. (2024). Menke-Hennekam syndrome; delineation of domain-specific subtypes with distinct clinical and DNA methylation profiles. HGG Adv. *5*, 100287. https://doi.org/10.1016/j.xhgg.2024.100287.

70. Pasqualucci, L., Dominguez-Sola, D., Chiarenza, A., Fabbri, G., Grunn, A., Trifonov, V., Kasper, L.H., Lerach, S., Tang, H., Ma, J., et al. (2011). Inactivating mutations of acetyltransferase genes in B-cell lymphoma. Nature *471*, 189–195.

71. Murata, T., Kurokawa, R., and Krones, A. (2001). Defect of histone acetyltransferase activity of the nuclear transcriptional coactivator CBP in Rubinstein–Taybi syndrome. Hum. Mol. Genet.

72. Kalkhoven, E., Roelfsema, J.H., Teunissen, H., den Boer, A., Ariyurek, Y., Zantema, A., Breuning, M.H., Hennekam, R.C.M., and Peters, D.J.M. (2003). Loss of CBP acetyltransferase activity by PHD finger mutations in Rubinstein-Taybi syndrome. Hum. Mol. Genet. *12*, 441–450.

73. Mensah, M.A., Niskanen, H., Magalhaes, A.P., Basu, S., Kircher, M., Sczakiel, H.L., Reiter, A.M.V., Elsner, J., Meinecke, P., Biskup, S., et al. (2023). Aberrant phase separation and nucleolar dysfunction in

rare genetic diseases. Nature, 1–8. https://doi.org/10.1038/s41586-022-05682-1.

74. Narita, T., Ito, S., Higashijima, Y., Chu, W.K., Neumann, K., Walter, J., Satpathy, S., Liebner, T., Hamilton, W.B., Maskey, E., et al. (2021). Enhancers are activated by p300/CBP activity-dependent PIC assembly, RNAPII recruitment, and pause release. Mol. Cell. https://doi.org/10.1016/j.molcel.2021.03.008.

75. Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308. https://doi.org/10.1038/nprot.2013.143.

76. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

77. Zhang, S., Cooper, J.A.L., Chong, Y.S., Naveed, A., Mayoh, C., Jayatilleke, N., Liu, T., Amos, S., Kobelke, S., Marshall, A.C., et al. (2023). NONO enhances mRNA processing of super-enhancer-associated GATA2 and HAND2 genes in neuroblastoma. EMBO Rep. *24*. https://doi.org/10.15252/embr.202254977.

78. R Core Team (2021). R: A Language and Environment for Statistical Computing. Preprint at R Foundation for Statistical Computing.

79. Thévenaz, P., Ruttimann, U.E., and Unser, M. (1998). A pyramid approach to subpixel registration based on intensity. IEEE Trans. Image Process. *7*, 27–41. https://doi.org/10.1109/83.650848.

80. San Diego, C.U. (2022). GraphPad Prism 9.0 Software.

81. Babraham bioinformatics - FastQC A quality control tool for high throughput sequence data http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

82. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal *17*, 10–12. https://doi.org/10.14806/ej.17.1.200.

83. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

84. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

85. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

86. Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S., and Karolchik, D. (2010). BigWig and BigBed: enabling browsing of large distributed datasets. Bioinformatics *26*, 2204–2207. https://doi.org/10.1093/bioinformatics/btq351.

87. Ewels, P.A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M.U., Di Tommaso, P., and Nahnsen, S. (2020). The nf-core framework for community-curated bioinformatics pipelines. Nat. Biotechnol. *38*, 276–278. https://doi.org/10.1038/s41587-020-0439-x.

88. Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44*, W160–W165. https://doi.org/10.1093/nar/gkw257.

89. Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

90. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for

assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

91. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

92. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

93. Jones, D.T., and Cozzetto, D. (2015). DISOPRED3: precise disordered region predictions with annotated protein-binding activity. Bioinformatics *31*, 857–863. https://doi.org/10.1093/bioinformatics/btu744.

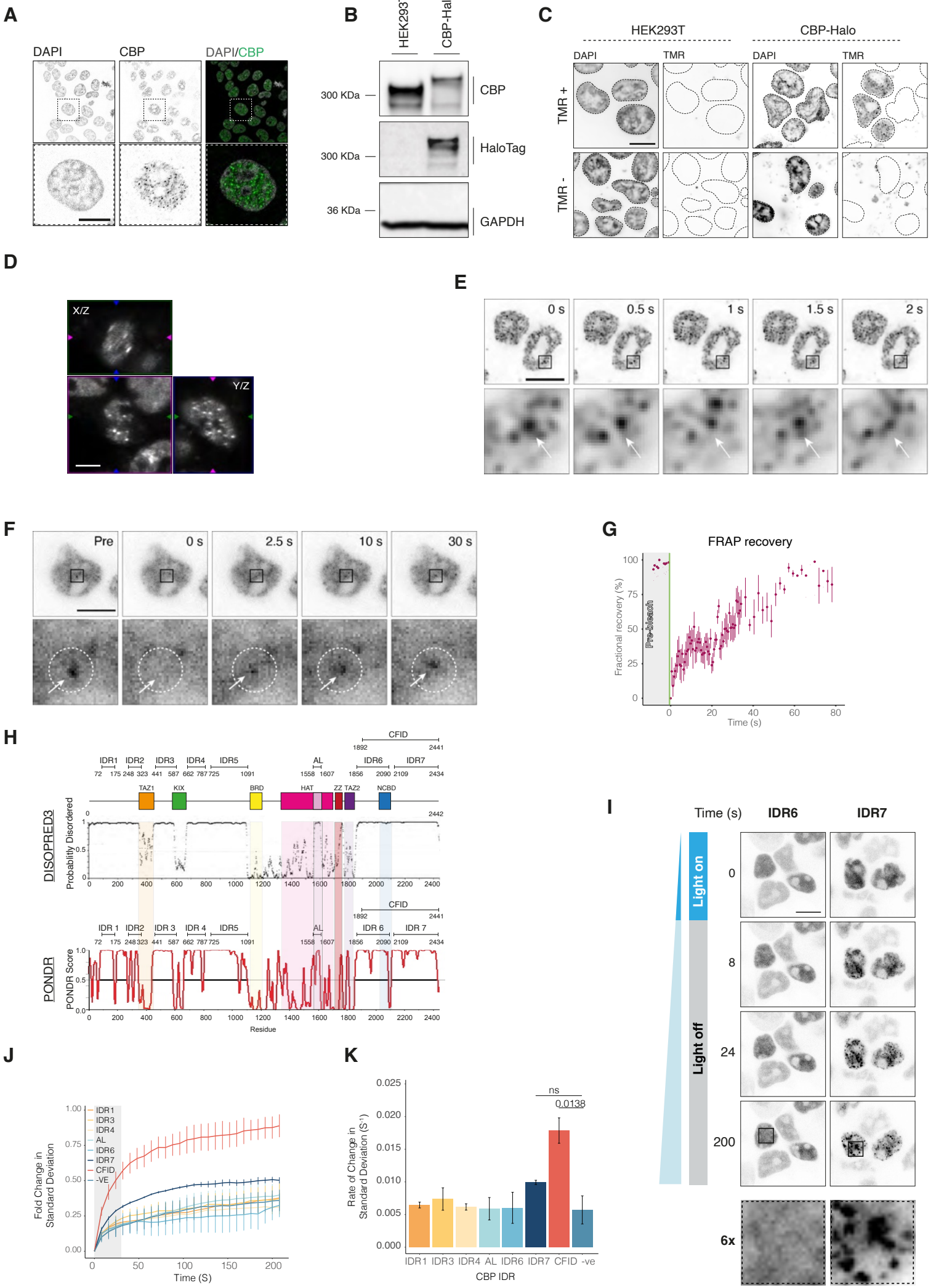
94. Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., et al. (2021). Highly accurate protein structure prediction for the human proteome. Nature *596*, 590–596. https://doi.org/10.1038/s41586-021-03828-1.

95. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci. *30*, 70–82. https://doi.org/10.1002/pro.3943.

96. Conant, D., Hsiau, T., Rossi, N., Oki, J., Maures, T., Waite, K., Yang, J., Joshi, S., Kelso, R., Holden, K., et al. (2022). Inference of CRISPR edits from Sanger trace data. CRISPR J. *5*, 123–130. https://doi.org/10.1089/crispr.2021.0113.

97. Madeira, F., Pearce, M., Tivey, A.R.N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., and Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Res. *50*, W276–W279. https://doi.org/10.1093/nar/gkac240.

Figure 1


A
DAPI
CBP
DAPI/CBP
B
HEK293T
CBP-Halo
300 KDa
CBP
300 KDa
HaloTag
36 KDa
GAPDH
C
HEK293T
CBP-Halo
DAPI
TMR
DAPI
TMR
TMR +
TMR -
D
X/Z
Y/Z
E
0 s
0.5 s
1 s
1.5 s
2 s
F
Pre
0 s
2.5 s
10 s
30 s
G
FRAP recovery
Fractional recovery (%)
Pre-bleach
Time (s)
H
CFID
IDR1
IDR2
IDR3
IDR4
IDR5
AL
IDR6
IDR7
TAZ1
KIX
BRD
HAT
ZZ
TAZ2
NCBD
DISOPRED3
Probability Disordered
PONDR
PONDR Score
Residue
I
Time (s)
IDR6
IDR7
Light on
Light off
0
8
24
200
6x
J
IDR1
IDR3
IDR4
AL
IDR6
IDR7
CFID
-VE
Fold Change in Standard Deviation
Time (S)
K
ns
0.0138
Rate of Change in Standard Deviation (S⁻¹)
IDR1
IDR3
IDR4
AL
IDR6
IDR7
CFID
-ve
CBP IDR

Supplementary Figure 1

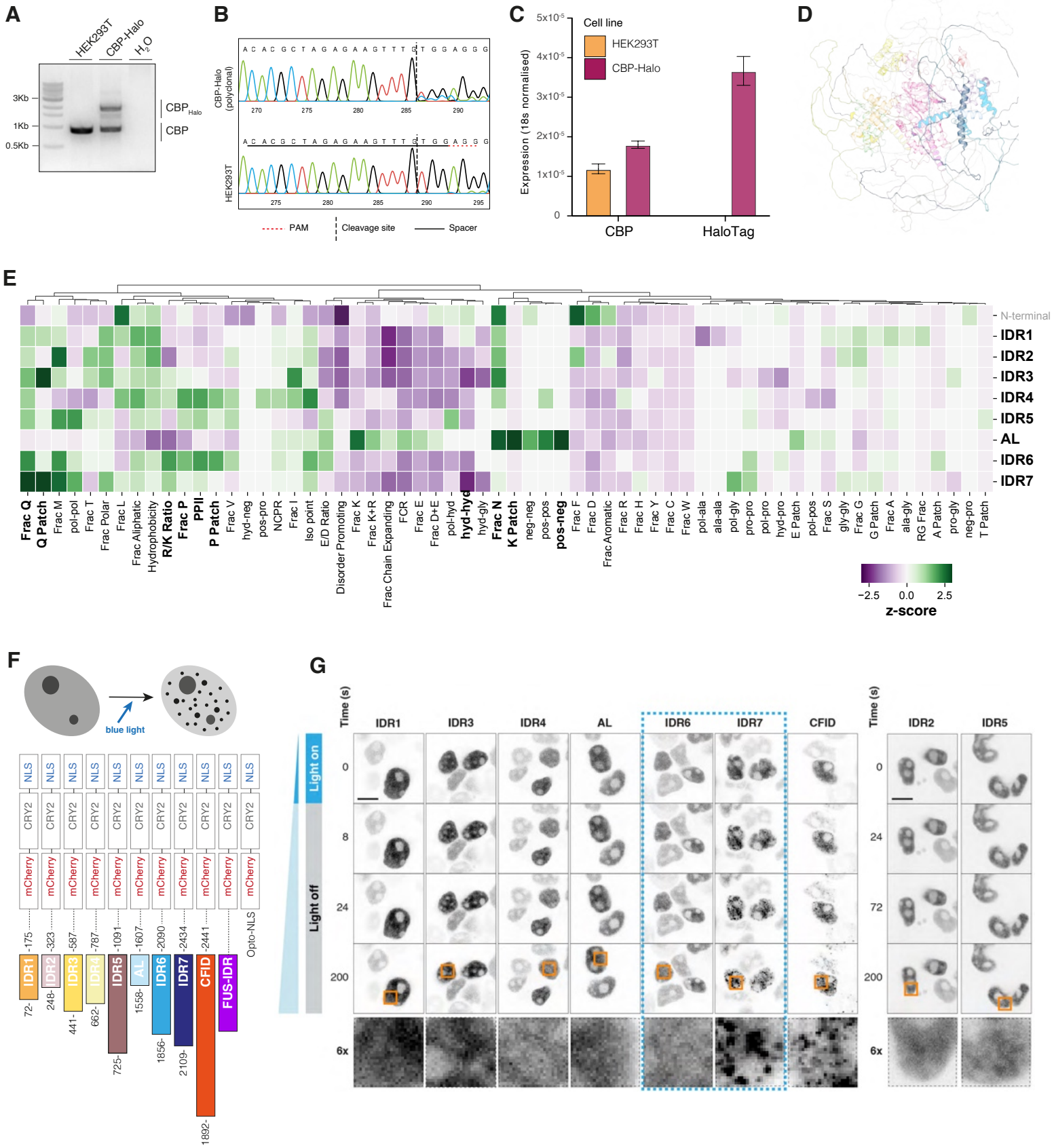
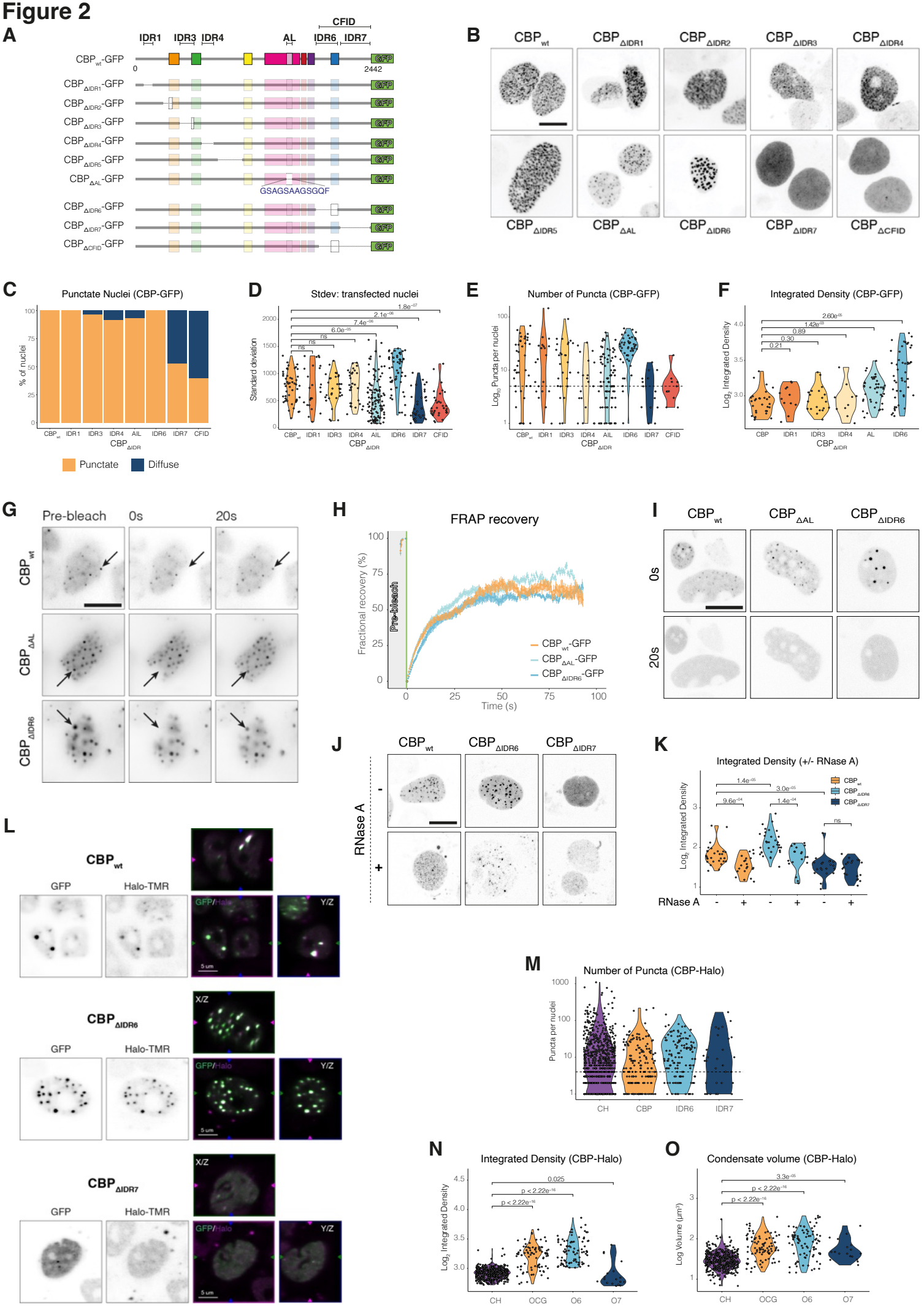
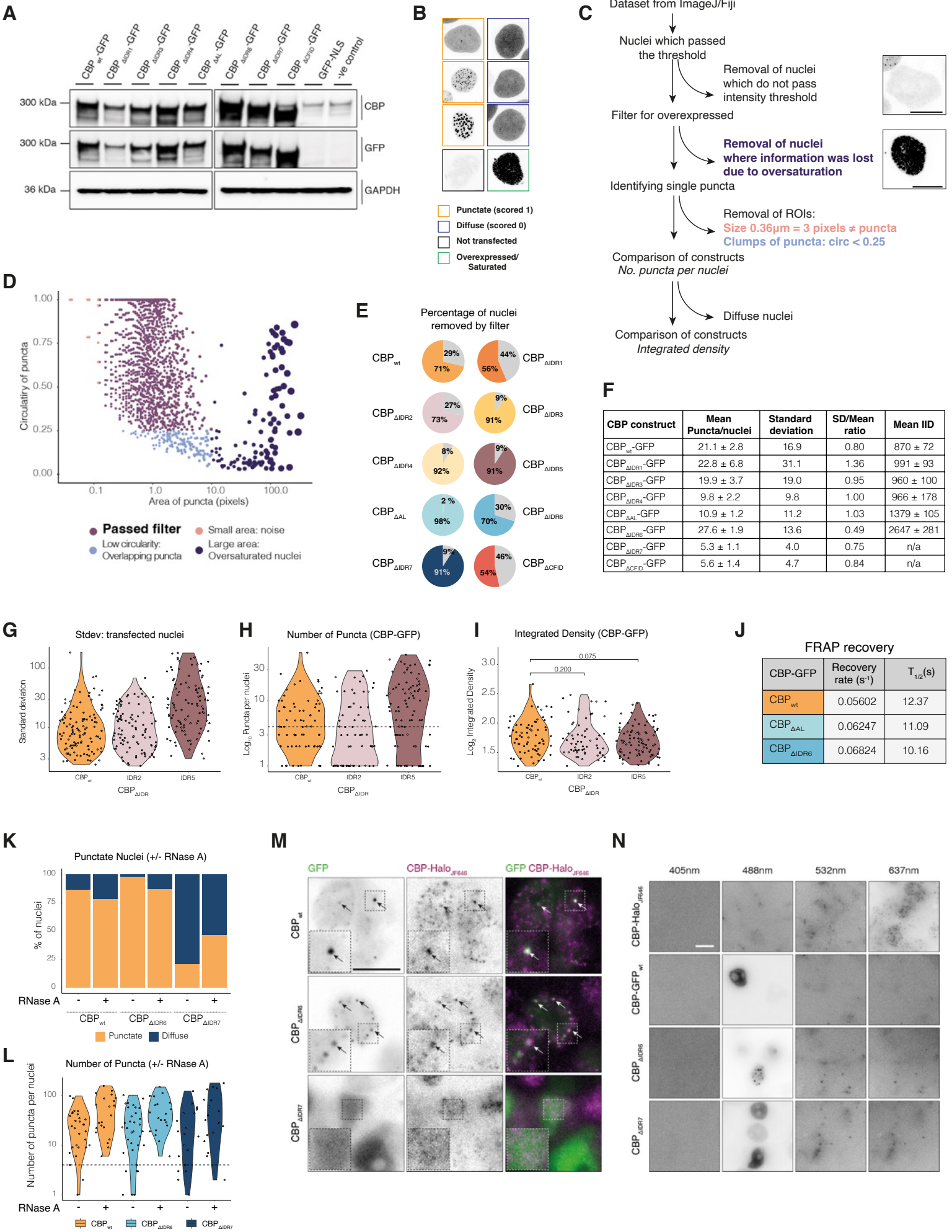


Figure 2


A
IDR1
IDR3
IDR4
AL
CFID
IDR6
IDR7
CBP$_{wt}$-GFP
0
2442
CBP$_{\Delta IDR1}$-GFP
CBP$_{\Delta IDR2}$-GFP
CBP$_{\Delta IDR3}$-GFP
CBP$_{\Delta IDR4}$-GFP
CBP$_{\Delta IDR5}$-GFP
CBP$_{\Delta AL}$-GFP
GSAGSAAGSGQF
CBP$_{\Delta IDR6}$-GFP
CBP$_{\Delta IDR7}$-GFP
CBP$_{\Delta CFID}$-GFP
GFP
B
CBP$_{wt}$
CBP$_{\Delta IDR1}$
CBP$_{\Delta IDR2}$
CBP$_{\Delta IDR3}$
CBP$_{\Delta IDR4}$
CBP$_{\Delta IDR5}$
CBP$_{\Delta AL}$
CBP$_{\Delta IDR6}$
CBP$_{\Delta IDR7}$
CBP$_{\Delta CFID}$
C
Punctate Nuclei (CBP-GFP)
% of nuclei
CBP$_{\Delta IDR}$
Punctate
Diffuse
D
Stdev: transfected nuclei
Standard deviation
E
Number of Puncta (CBP-GFP)
Log$_{10}$ Puncta per nuclei
F
Integrated Density (CBP-GFP)
Log$_2$ Integrated Density
G
Pre-bleach
0s
20s
H
FRAP recovery
Fractional recovery (%)
Pre-bleach
Time (s)
CBP$_{wt}$-GFP
CBP$_{\Delta AL}$-GFP
CBP$_{\Delta IDR6}$-GFP
I
J
RNase A
K
Integrated Density (+/- RNase A)
L
GFP
Halo-TMR
GFP/Halo
X/Z
Y/Z
5 um
M
Number of Puncta (CBP-Halo)
Puncta per nuclei
CH
CBP
IDR6
IDR7
N
Integrated Density (CBP-Halo)
Log$_2$ Integrated Density
OCG
O6
O7
O
Condensate volume (CBP-Halo)
Log Volume (μm³)

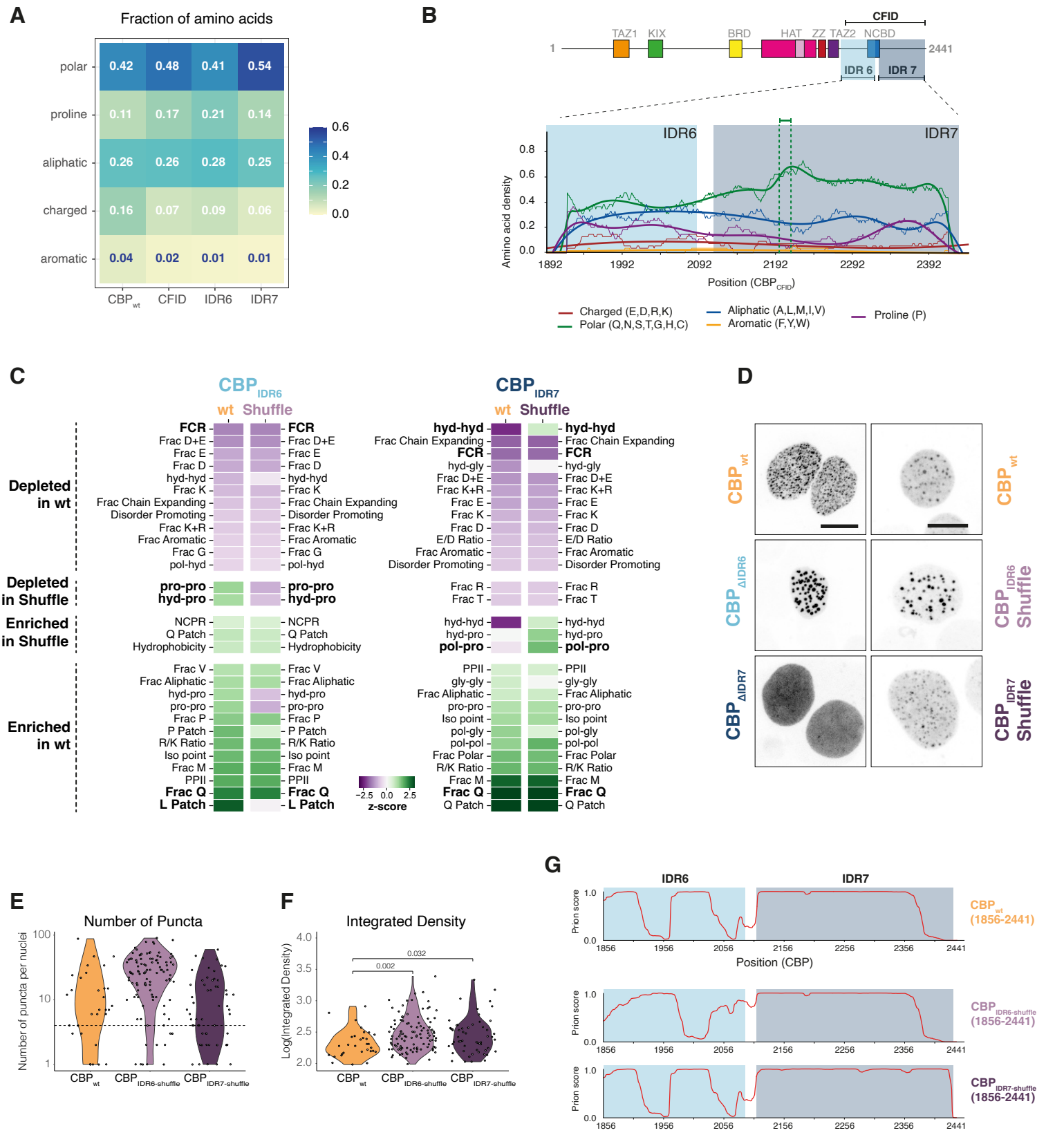
Supplementary Figure 2



CBP construct	Mean Puncta/nuclei	Standard deviation	SD/Mean ratio	Mean IID
CBP_{wt}-GFP	21.1 ± 2.8	16.9	0.80	870 ± 72
$CBP_{\Delta IDR1}$-GFP	22.8 ± 6.8	31.1	1.36	991 ± 93
$CBP_{\Delta IDR3}$-GFP	19.9 ± 3.7	19.0	0.95	960 ± 100
$CBP_{\Delta IDR4}$-GFP	9.8 ± 2.2	9.8	1.00	966 ± 178
$CBP_{\Delta AL}$-GFP	10.9 ± 1.2	11.2	1.03	1379 ± 105
$CBP_{\Delta IDR6}$-GFP	27.6 ± 1.9	13.6	0.49	2647 ± 281
$CBP_{\Delta IDR7}$-GFP	5.3 ± 1.1	4.0	0.75	n/a
$CBP_{\Delta CFID}$-GFP	5.6 ± 1.4	4.7	0.84	n/a

CBP-GFP	Recovery rate (s^{-1})	$T_{1/2}$(s)
CBP_{wt}	0.05602	12.37
$CBP_{\Delta AL}$	0.06247	11.09
$CBP_{\Delta IDR6}$	0.06824	10.16

Figure 3


A
Fraction of amino acids
polar
proline
aliphatic
charged
aromatic
0.42 0.48 0.41 0.54
0.11 0.17 0.21 0.14
0.26 0.26 0.28 0.25
0.16 0.07 0.09 0.06
0.04 0.02 0.01 0.01
CBP wt
CFID
IDR6
IDR7
0.6
0.4
0.2
0.0
B
CFID
TAZ1
KIX
BRD
HAT
ZZ
TAZ2
NCBD
1
2441
IDR 6
IDR 7
IDR6
IDR7
Amino acid density
Position (CBP CFID)
Charged (E,D,R,K)
Polar (Q,N,S,T,G,H,C)
Aliphatic (A,L,M,I,V)
Aromatic (F,Y,W)
Proline (P)
C
CBP IDR6
wt Shuffle
CBP IDR7
wt Shuffle
Depleted in wt
Depleted in Shuffle
Enriched in Shuffle
Enriched in wt
z-score
D
CBP wt
CBP ΔIDR6
CBP ΔIDR7
CBP IDR6 Shuffle
CBP IDR7 Shuffle
E
Number of Puncta
Number of puncta per nuclei
CBP wt
CBP IDR6-shuffle
CBP IDR7-shuffle
F
Integrated Density
Log(Integrated Density)
0.032
0.002
G
IDR6
IDR7
Prion score
Position (CBP)
CBP wt (1856-2441)
CBP IDR6-shuffle (1856-2441)
CBP IDR7-shuffle (1856-2441)

Supplementary Figure 3

A

IDR6 Shuffle (residues 1856-2090)

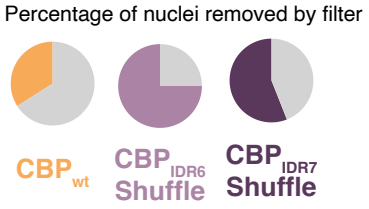
PGLPSPVQNLLPSPSMQSIRAQVATQPQPH-
QVSEGPAAAPRMLNPLQQRGPHPAAMVP-
PVPGSPQPTPGSPQADPRSPLTPPQQMQPTLTL
QQNSGQMQAEVSGPMQAPVTQPPQPIQQMPSRQ
RPRSAATNSFTWGPPVGTIVPAQMQRMMVAPPP
RGSKPAIPSLQERQRPVTRNRPQSSYSSSPVTT
PQSKQANTMLPRNQANMPSIVVVARTLQQPTLQ
SQQAPALGPQPP

B

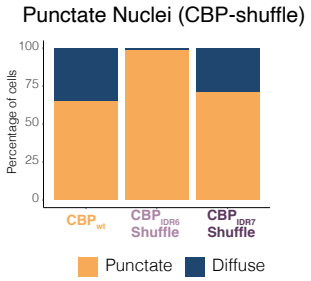
IDR7 Shuffle (residues 2109-2434)

FQPQGDNVQPPQGSLAQQYEQSQLPQAMDMP-
GLMGQSSGPPPTETPGAPPAPQGRDGPAS-
QQMQPLQSARPPAVSQQRMMNTQQMGSQGQQGV
HQLMRGGRQQGQQPGLNQVMAHALPQMHMLQIQ
LLPQSTQQSGMQESQQIGQQPQQMQQHASHKNG
QGQSQSSQNINSSSIRSLLPAPSTQPPISMPGA
AANPPPQLANSLQHISQQTPQQVQYGPRQGAPI
VGPPQMSLQLPQLQAQQLPHGSAHGLLGRHSLT
MQPMGGMLQPQRQNQQPQGQLAQRQVMQNSVQG
HSGAMMMGGPAPPQNPQNPGDLQTSSGQNQPHP
GN

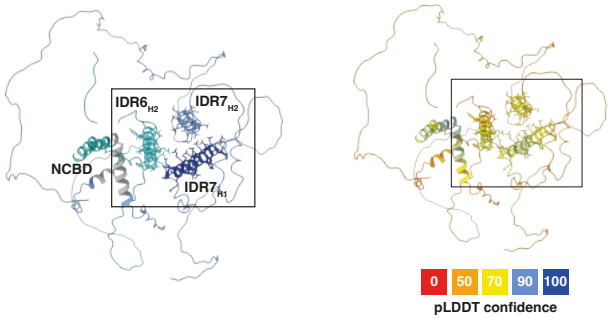
C



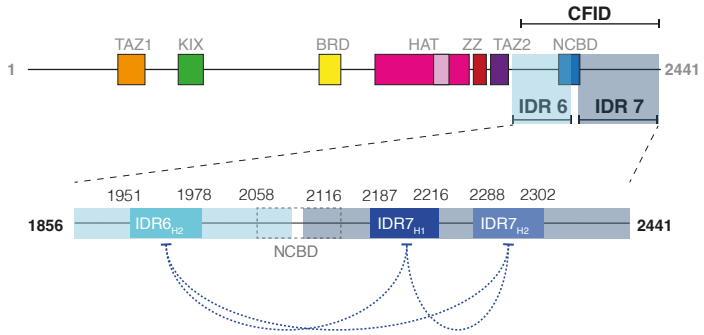
D



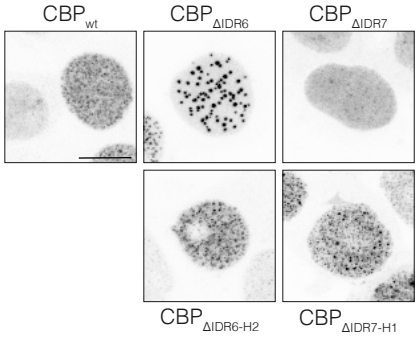
E



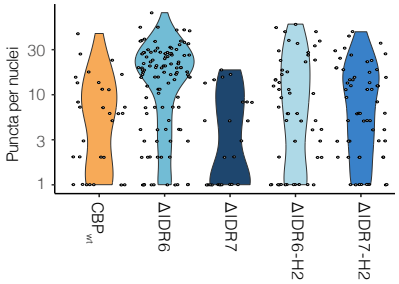
F



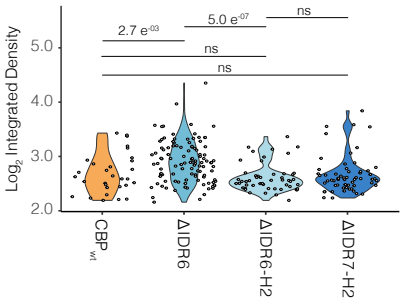
G



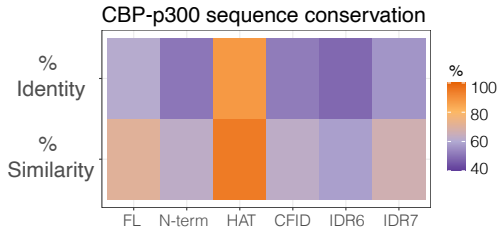
H



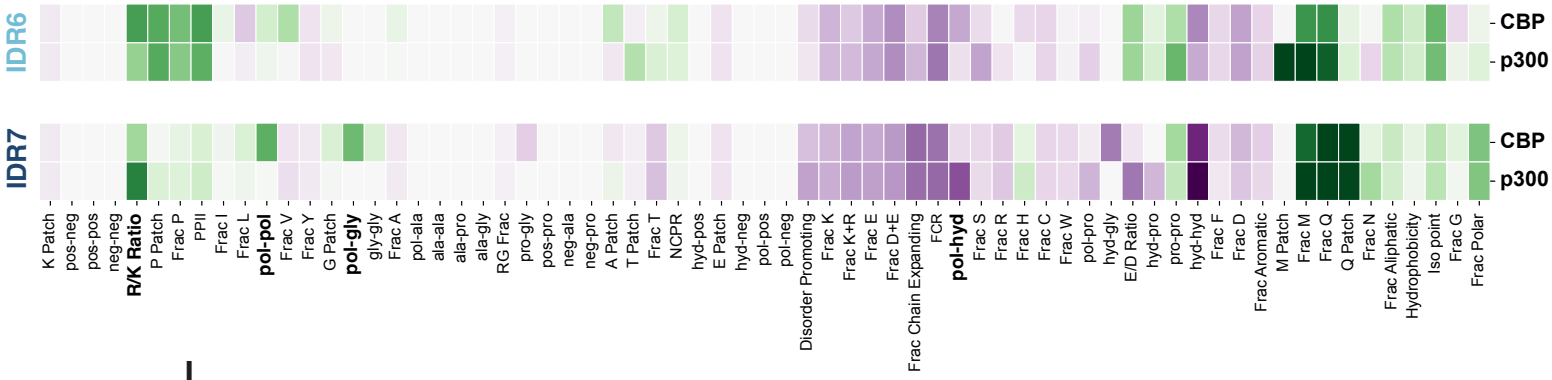
I



J



K



L

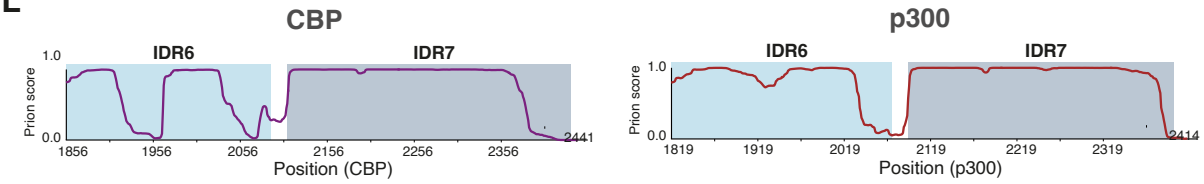
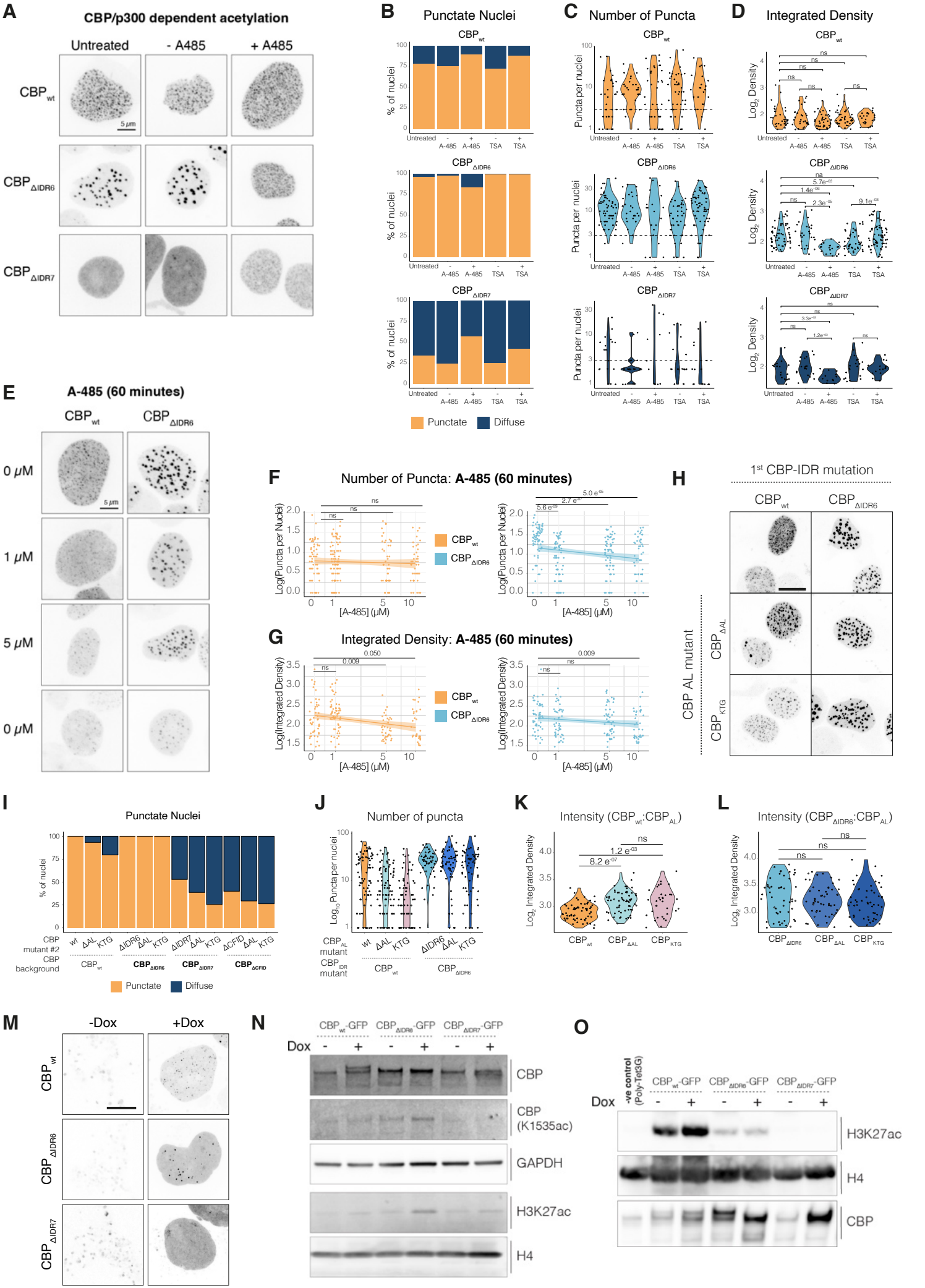


Figure 4


A
CBP/p300 dependent acetylation
Untreated
- A485
+ A485
CBP wt
CBP ΔIDR6
CBP ΔIDR7
5 µm
B
Punctate Nuclei
% of nuclei
Untreated
A-485
TSA
C
Number of Puncta
Puncta per nuclei
D
Integrated Density
Log2 Density
ns
Punctate
Diffuse
E
A-485 (60 minutes)
0 µM
1 µM
5 µM
F
Number of Puncta: A-485 (60 minutes)
Log(Puncta per Nuclei)
[A-485] (µM)
G
Integrated Density: A-485 (60 minutes)
Log(Integrated Density)
H
1st CBP-IDR mutation
CBP AL mutant
CBP ΔAL
CBP KTG
I
Punctate Nuclei
CBP mutant #2
CBP background
J
Number of puncta
Log10 Puncta per nuclei
CBP AL mutant
CBP IDR mutant
K
Intensity (CBP wt:CBP AL)
Log2 Integrated Density
L
Intensity (CBP ΔIDR6:CBP AL)
M
-Dox
+Dox
N
CBP wt-GFP
CBP ΔIDR6-GFP
CBP ΔIDR7-GFP
Dox
CBP
CBP (K1535ac)
GAPDH
H3K27ac
H4
O
-ve control (Poly-Tet3G)
H3K27ac
H4
CBP

Supplementary Figure 4

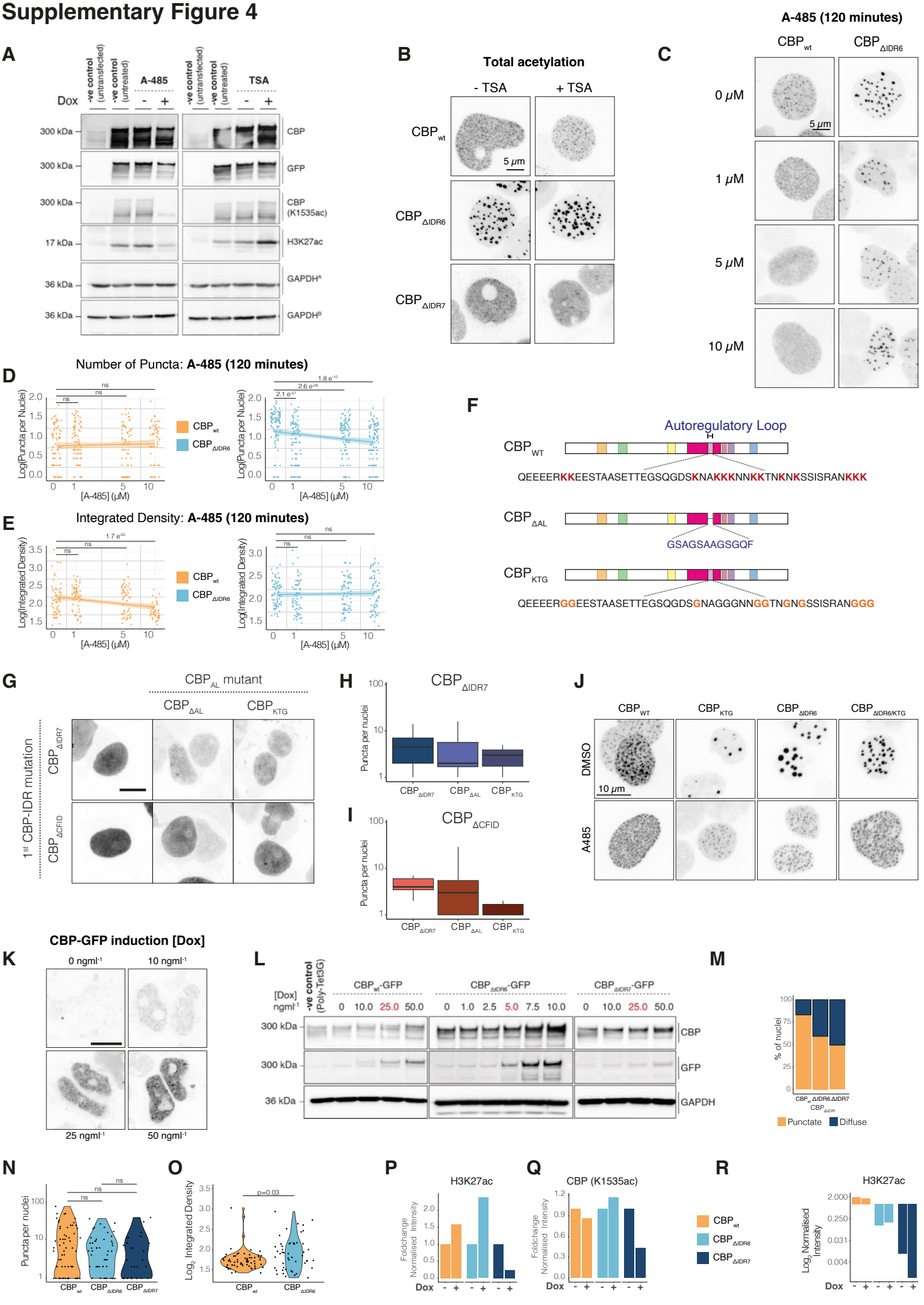
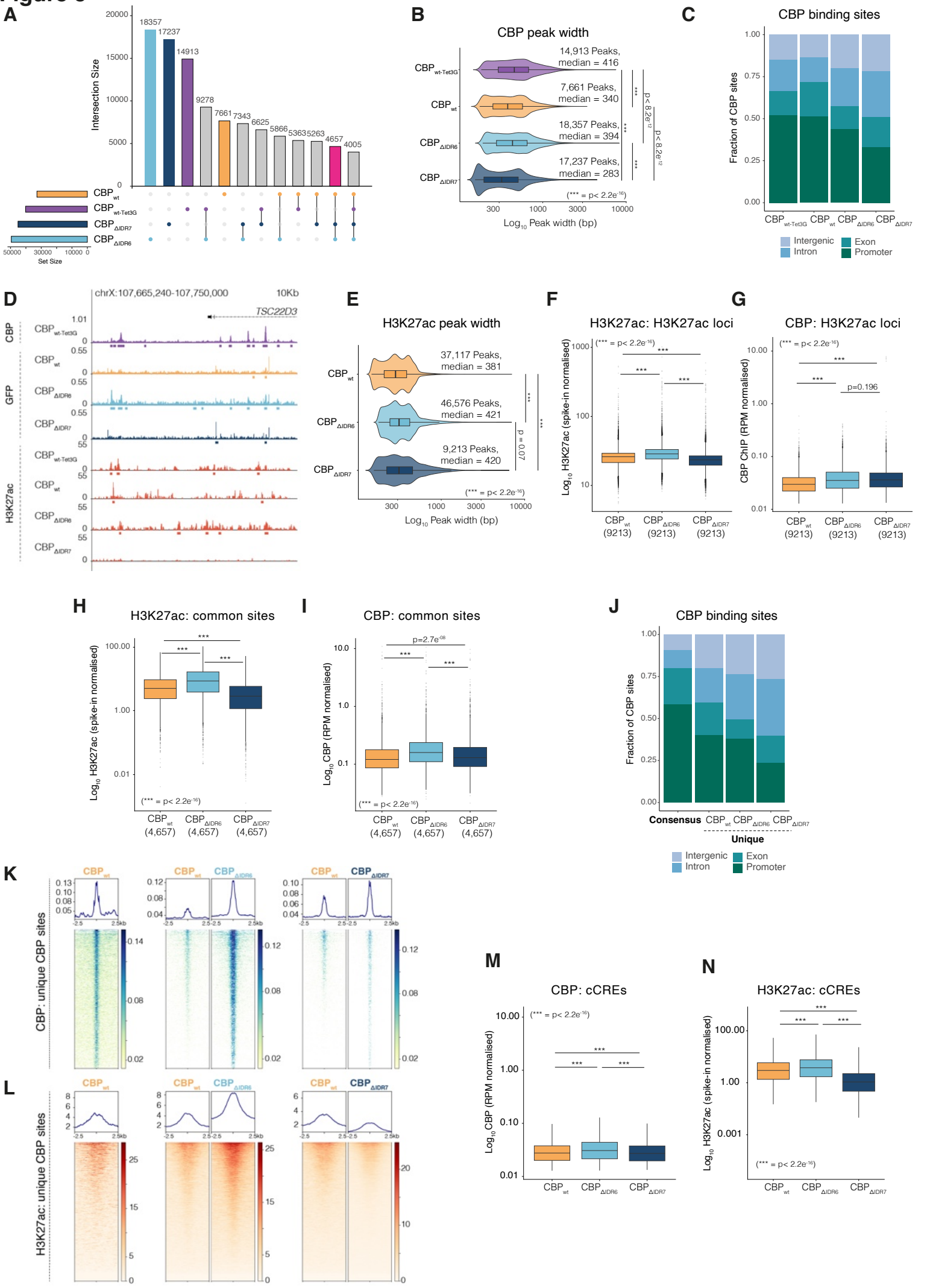


Figure 5


A
Intersection Size
18357
17237
14913
9278
7661
7343
6625
5866
5363
5263
4657
4005
CBP wt
CBP wt-Tet3G
CBP ΔIDR7
CBP ΔIDR6
Set Size
B
CBP peak width
14,913 Peaks, median = 416
7,661 Peaks, median = 340
18,357 Peaks, median = 394
17,237 Peaks, median = 283
(*** = p< 2.2e-16)
Log10 Peak width (bp)
C
CBP binding sites
Fraction of CBP sites
Intergenic
Intron
Exon
Promoter
D
chrX:107,665,240-107,750,000
10Kb
TSC22D3
CBP
GFP
H3K27ac
E
H3K27ac peak width
37,117 Peaks, median = 381
46,576 Peaks, median = 421
9,213 Peaks, median = 420
p = 0.07
(*** = p< 2.2e-16)
Log10 Peak width (bp)
F
H3K27ac: H3K27ac loci
Log10 H3K27ac (spike-in normalised)
(9213)
G
CBP: H3K27ac loci
CBP ChIP (RPM normalised)
p=0.196
H
H3K27ac: common sites
Log10 H3K27ac (spike-in normalised)
(4,657)
I
CBP: common sites
p=2.7e-08
Log10 CBP (RPM normalised)
J
CBP binding sites
Consensus
Unique
K
CBP: unique CBP sites
L
H3K27ac: unique CBP sites
M
CBP: cCREs
Log10 CBP (RPM normalised)
N
H3K27ac: cCREs
Log10 H3K27ac (spike-in normalised)

Supplementary Figure 5

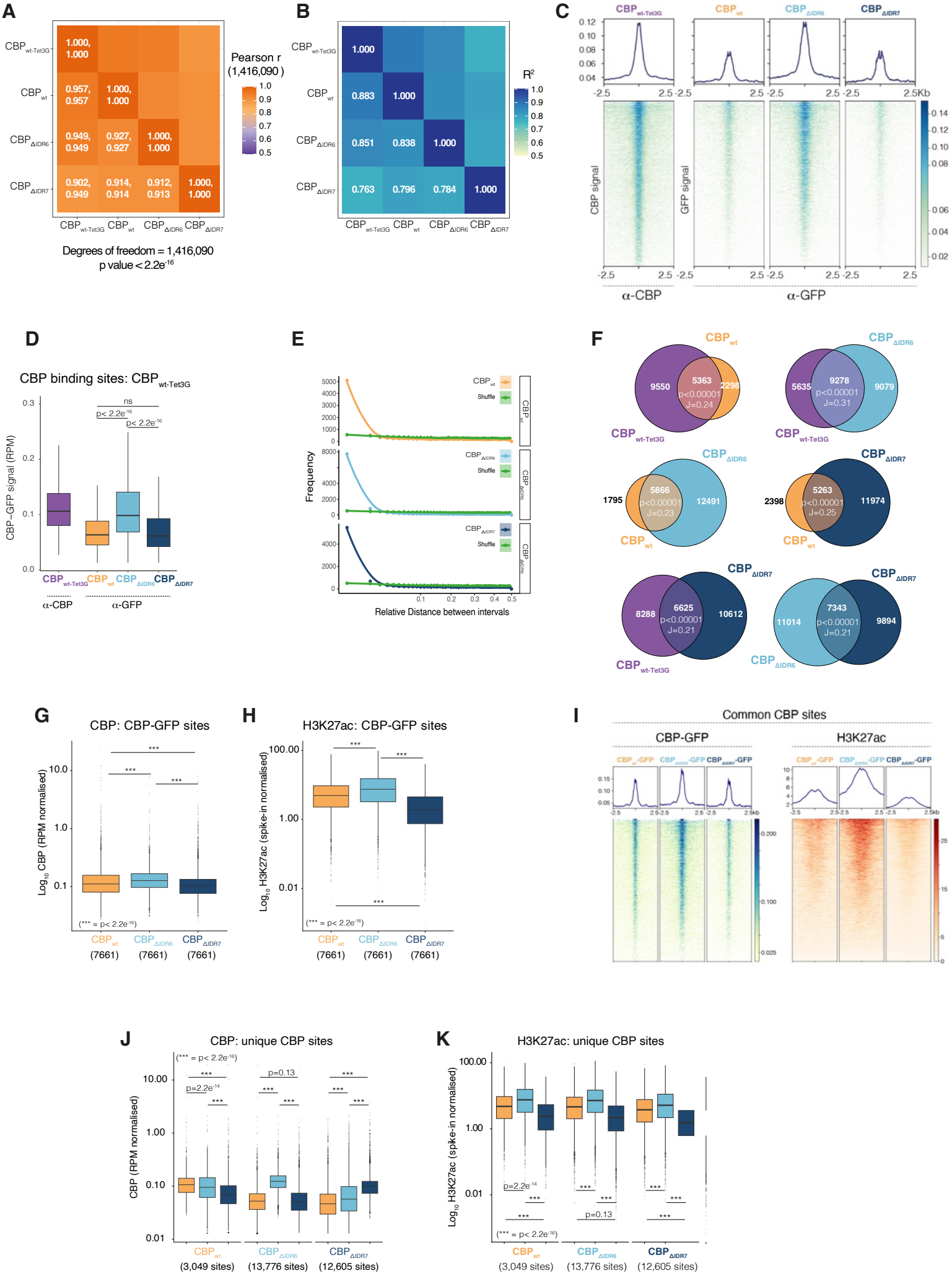
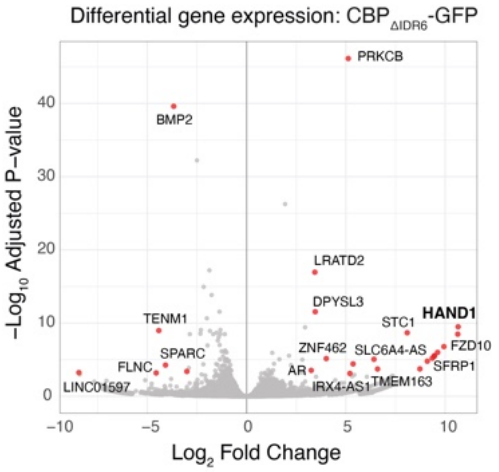
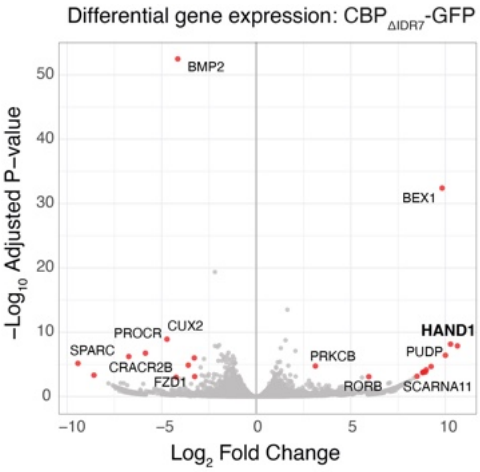


Figure 6

A

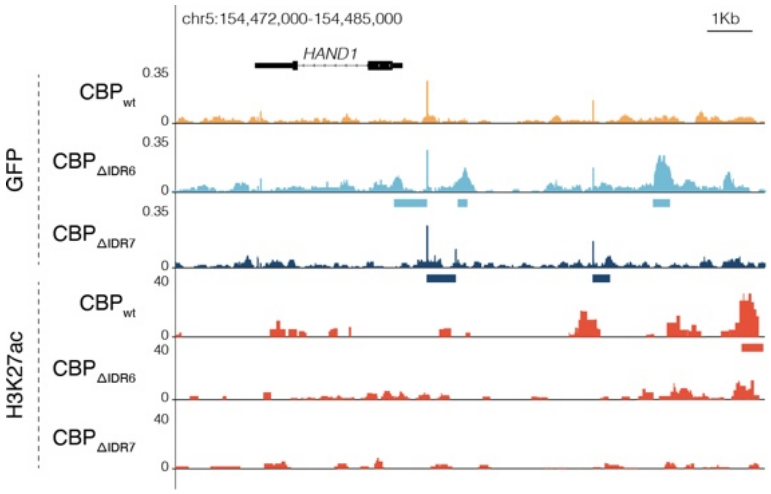


B



Not Significant • Significant

C



Supplementary Figure 6

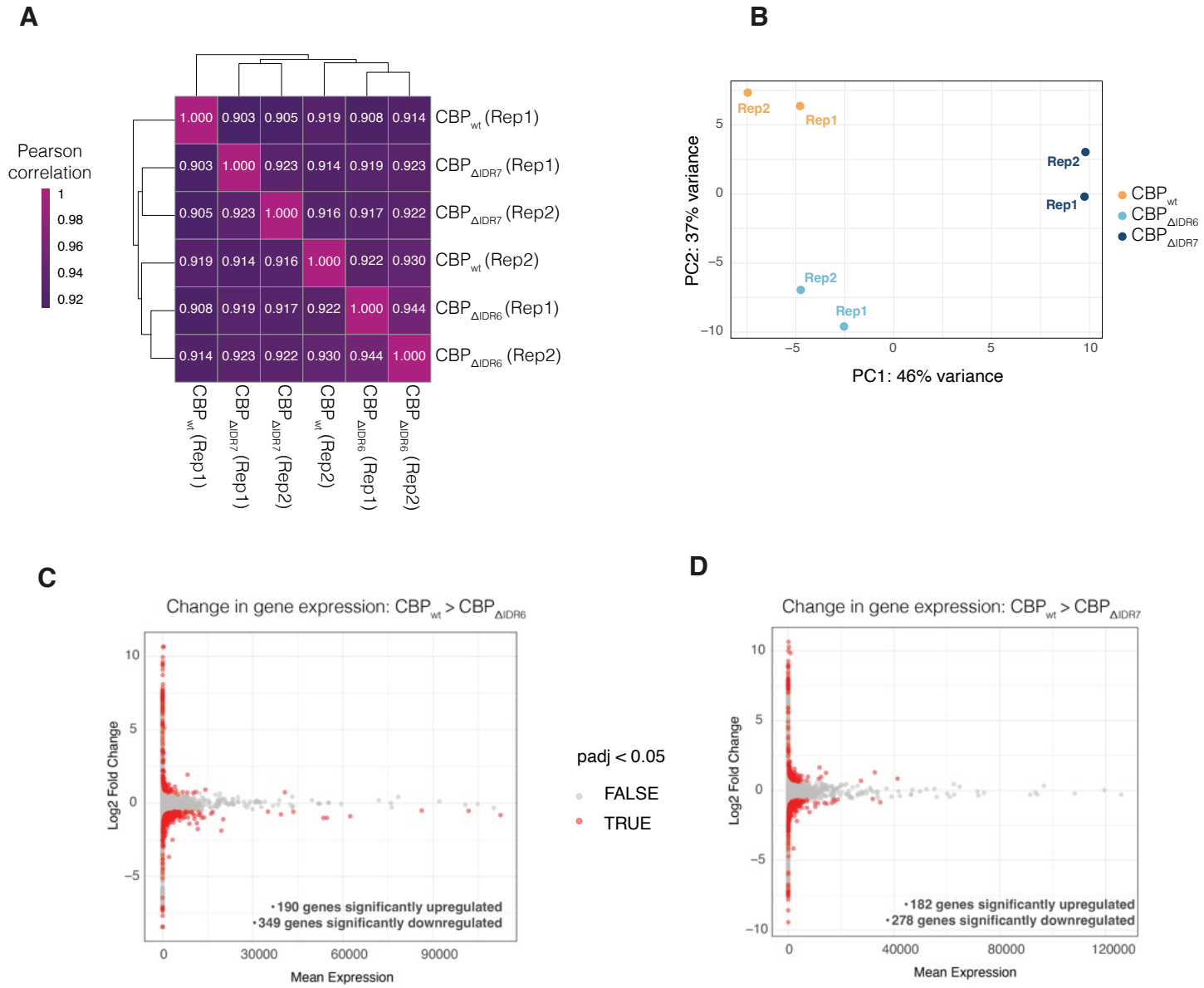
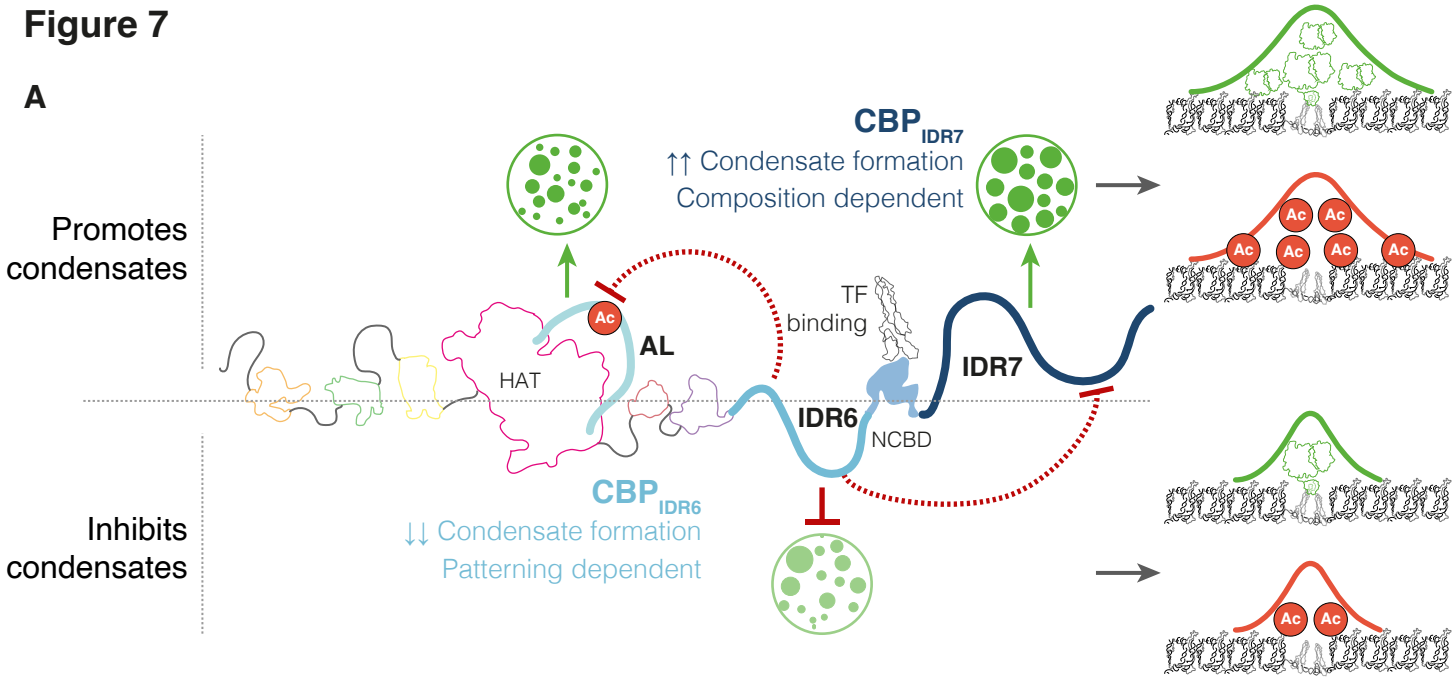
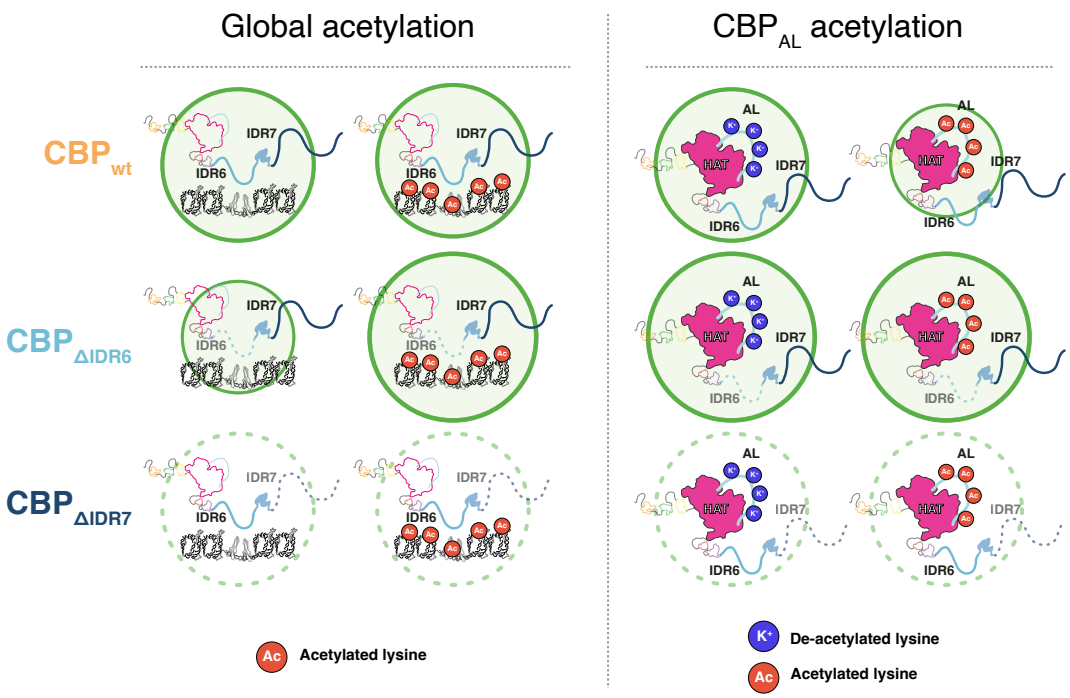
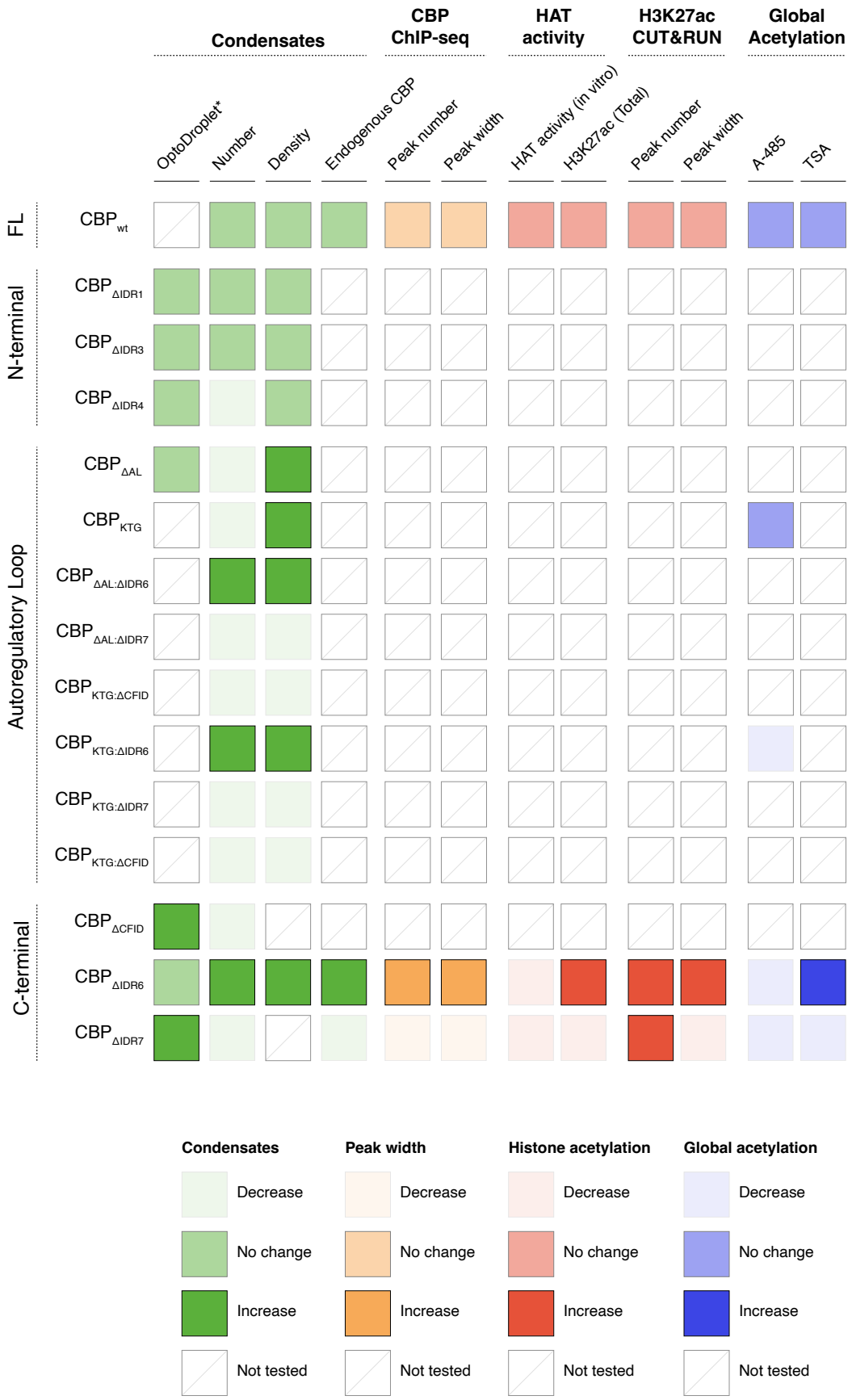


Figure 7

A


Promotes condensates
Inhibits condensates
CBP IDR7
↑↑ Condensate formation
Composition dependent
CBP IDR6
↓↓ Condensate formation
Patterning dependent
HAT
AL
Ac
TF binding
IDR7
IDR6
NCBD

B


Global acetylation
CBP AL acetylation
CBP wt
CBP ΔIDR6
CBP ΔIDR7
IDR7
IDR6
AL
HAT
Acetylated lysine
De-acetylated lysine

Supplementary Figure 7